1/54

#	1 mM GSH	0.1 mM GSSG	0.3 mM LM	mM NaCl	mM KCl	0.055% PEG 3350	550 mM GndHCl	1.1 mM EDTA	2.2 mM MgCl$_2$	2.2 mM CaCl$_2$	440 mM Sucrose	550 mM L-Arg	Activity U/g IB
2 (55 mM MES pH 6.5)	+	+	+	10.56	0.44	0	+	0	+	+	0	0	0
3 (55 mM MES pH 6.5)	+	+	0	10.56	0.44	+	+	+	0	0	+	+	0
#5 (55 mM MES pH 6.5)	+	+	0	264	11	0	0	0	+	+	+	0	0
#8 (55 mM MES pH 6.5)	+	+	+	264	11	+	0	+	0	0	0	+	40.00
#10 (55 mM MES pH 8.2)	+	+	+	10.56	0.44	0	0	+	0	0	+	0	0
#11 (55 mM Tris pH 8.2)	+	+	0	10.56	0.44	+	0	0	+	+	0	+	105.26
#13 (55 mM Tris pH 8.2)	+	+	0	264	11	0	+	+	0	0	0	0	15.65
#16 (55 mM Tris pH 8.2)	+	+	+	264	11	+	+	0	+	+	+	+	48.70

FIG. 1

ELution of MBP-ST3Gal III from Amylose column
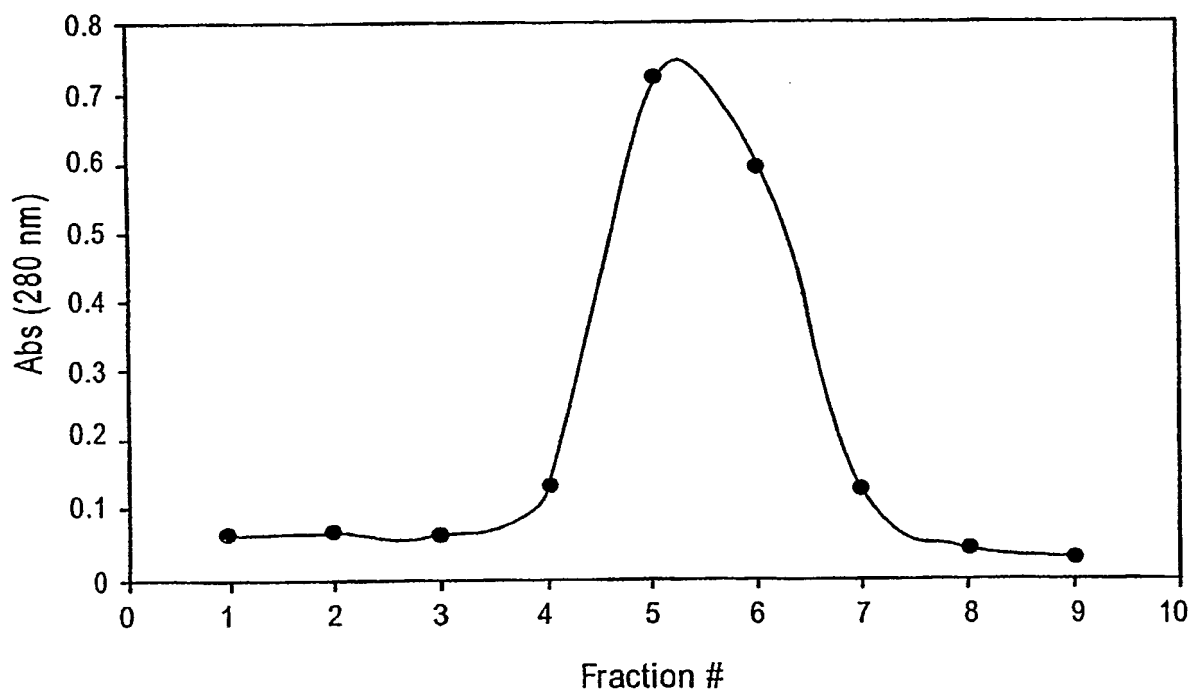
FIG. 2

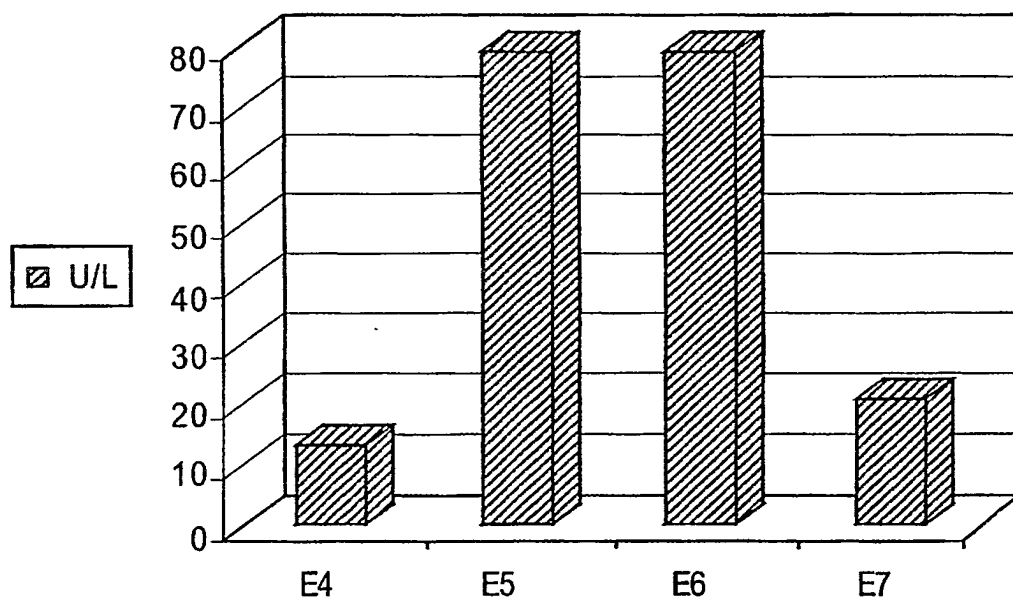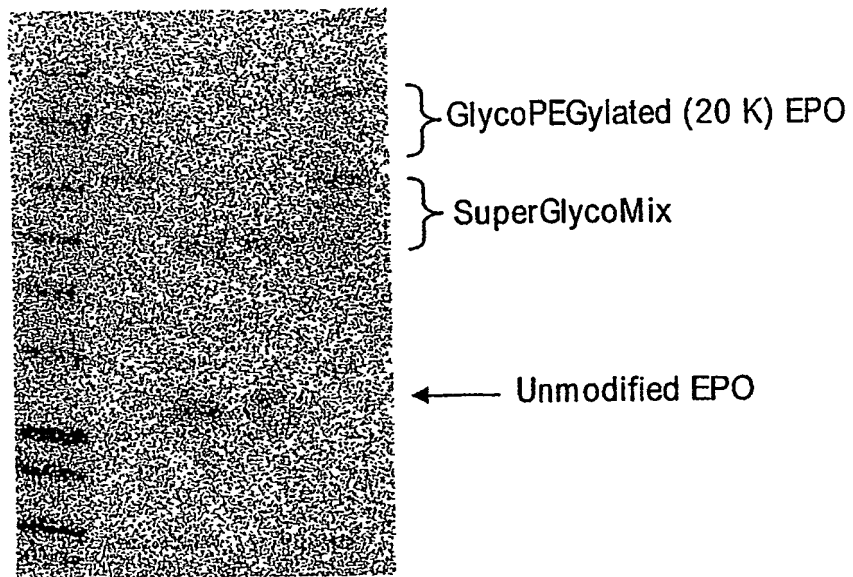**ST3 Gal III activities of the Amylose purified refolded MBP-ST3Gal III fractions**

FIG. 3

20 K
Glyco-
PEGylated   10 K
transferrins   GlycoPEGylated
                   tranferrins

FIG. 4

GlycoPEGylation (20 K) of EPO

FIG. 5

6/54

```
           10        20        30        40        50        60
/usr/t MLKKQSAGLVLWGAILFVAWNALLLLFFWTRPAPGRPPSVSALDGDPASLTREVIRLAQD
       ::::::::::::::::::::::::::::::::::::::::..: :: .::: :::::::::::::
P27115 MLKKQSAGLVLWGAILFVAWNALLLLFFWTRPVPSRLPSDNALDDDPASLTREVIRLAQD
           10        20        30        40        50        60


           70        80        90       100       110
/usr/t AEVELERQRGLLQQIGD--ALSSQRGRVPTAAPPAQPRVPVTPAPAVIPILVIACDRSTV
       :::::::::::::::: .   :: :::  .:::::::::::.::::: :::::::::::::::
P27115 AEVELERQRGLLQQIREHHALWSQRWKVPTAAPPAQPHVPVTPPPAVIPILVIACDRSTV
           70        80        90       100       110       120


         120       130       140       150       160       170
/usr/t RRCLDKLLHYRPSAELFPIIVSQDCGHEETAQAIASYGSAVTHIRQPDLSSIAVPPDHRK
       :::::::::::::::::::::::::::::::::::.:::::::::::::::::.::: :::::
P27115 RRCLDKLLHYRPSAELFPIIVSQDCGHEETAQVIASYGSAVTHIRQPDLSNIAVQPDHRK
         130       140       150       160       170       180


         180       190       200       210       220       230
/usr/t FQGYYKIARHYRWAIGQVFRQFRFPAAVVVEDDLEVAPDFFEYFRATYPLLKADPSLWCV
       :::::::::::::::::::..: .:::::::::::::::::::::::.::::::::::::::
P27115 FQGYYKIARHYRWAIGQIFHNFNYPAAVVVEDDLEVAPDFFEYFQATYPLLKADPSLWCV
         190       200       210       220       230       240


         240       250       260       270       280       290
/usr/t SAWNDNGKEQMVDASRPELLYRTDFFPGLGWLLLAELWAELEPKWPKAFWDDWMRRPEQR
       :::::::::::::::..:.::::::::::::::::::::::::::::::::::::::::::::
P27115 SAWNDNGKEQMVDSSKPELLYRTDFFPGLGWLLLAELWAELEPKWPKAFWDDWMRRPEQR
         250       260       270       280       290       300


         300       310       320       330       340       350
/usr/t QGRACIRPEISRTMTFGRKGVSHGQFFDQHLKFIKLNQQFVHFTQLDLSYLQREAYDRDF
       .::::.::::::::::::::::::::::::::::::::::::::: :::::::::.:::::::
P27115 KGRACVRPEISRTMTFGRKGVSHGQFFDQHLKFIKLNQQFVPFTQLDLSYLQQEAYDRDF
         310       320       330       340       350       360


         360       370       380       390       400       410
/usr/t LARVYGAPQLQVEKVRTNDRKELGEVRVQYTGRDSFKAFAKALGVMDDLKSGVPRAGYRG
       ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
P27115 LARVYGAPQLQVEKVRTNDRKELGEVRVQYTGRDSFKAFAKALGVMDDLKSGVPRAGYRG
         370       380       390       400       410       420


         420       430       440
/usr/t IVTFQFRGRRVHLAPPPTWEGYDPSWN
       :::: ::::::::::: ::.:.:::::.
P27115 IVTFLFRGRRVHLAPPQTWDGYDPSWT
         430       440
```

FIG. 6

GnT1 Cys121Ser mutant

avipilviacdrstvrrsldkllhyrpsaelfpiivsqdcgheetaqaiasygsavthirqpdlssiavppdhrkfqgyykiarhyrwa
lgqvfrqfrfpaavvveddlevapdffeyfratypllkadpslwcvsawndngkeqmvdasrpellyrtdffpglgwlllaelwae
lepkwpkafwddwmrrpeqrqgracirpeisrtmtfgrkgvshgqffdqhlkfiknqqfvhftqldlsylqreaydrdflarvyg
apqlqvekvrtndrkelgevrvqytgrdsfkafakalgvmddlksgvpragyrgivtfqfpgrrvhlappptwegydpswn*

Gcggtgattcccatcctggtcatcgctgctgaccgcagcactgttcggcgcgctctctagacaagctgctgcattatcggccctcggctga
gctcttccccatcatcgttagccaggacgtgcggcacgagacgaccgcccagccatcgtcctctacggcagcgcgtcacgcaca
tccggcagcgcccgacctgagcgagccattgcggtgccgccgccgaccaccgcaagttccaggctactacaagatcgcgccactaccg
ctgggcgctgggcccaggtctccggcagtttcgcttcccgcgccgtggtcgttggtggaggatgacctgagggtggccccggacttctt
cgagtactttcggcgcccaccatccgcctgctgaaggccctcccctgtgtcgctcgcctgggctggtggtcgtgctgggctctggct
gatggtggacgccagcaggcctgagctgctctaccgcaccgcaccgactttttccctgccgggcggccgccgagcgcgccgctgcatacg
gagctggagccaagtggccaaaggccctttggccgcaaagatgacctttggccgcaagggtgagccacgggcacggggcggccgctgctgcagttatcaagctga
accagcagtttgtgcacttcacccagctgcacctgtcttacctgccatgaccggaggagccacgagatttcctcgcccgtcgtacggt
gctcccccagctgcaggtggagaaagtgaggaccaataagtgaccgaaggaggtctgggggtcggtgcggtgcagtatacgggcaggga
cagcttcaaggctttcgccaaggctctggtgatgatgacctaagtcggggttccgagagctgctaccggggtgattgtcacctt
ccagttccggccgcgccgtgtcacctggccgcgccccacccacgcgtggggaggggctatgatcctagctgtggaattag

FIG. 7

GnT1 Cys121Asp

avipilviacdrstvrrdldkllhyrpsaelfpiivsqdcgheetaqaiasygsavthirqpdlssiavppdhrkfqgyykiarhyrwa
lgqvfrqfrfpaavvveddlevapdffeyfratypllkadpslwcvsawndngkeqmvdasrpellyrtdffpglgwlllaelwae
lepkwpkafwddwmrrpeqrqgracirpeisrtmtfgrkgvshgqffdqhlkfiklnqqfvhftqldlsylqreaydrdflarvyg
apqlqvekvrtndrkelgevrvqytgrdsfkafakalgvmddlkssgvpragyrgivtfqfpgrrvhlappptwegydpswn*

Gcggtgattccatcctggtcatcgcctgtgaccgcagcactgttcggcgcgcgatcagacaagctgctgcattatcggccctcggctg
agctcttcccatcatcgttagccaggactgcgggcacgagaggacggcccagccatcgcctcctacgcagcgcggtcacgcac
atccggcagcccgacctgagcagcagcattgcggtgccgccgaccacaccgacctacacaagatcgcgccgccactacc
gctgggcgctgggcggcagttcttccggcagtttgcttcccccgcggcagtggtggtgaggatgacctgagggtggccccggacttct
tcgagtactttcgggccaccatccgtgctgaaggccgaccctcctgtggctcgtctcgcctggcgctgcttgttggccgagctctgggc
agatggtggacgccagagcgccagcagcctgctctaccgcctcctgggacgctgtctgctgctcggacgctggatgcgggccgcaggccctgcatac
tgagctggagccaagtggccaaaggccttgctggacgactttggccgcaaggggttgagcagcgcggagcacgtggagccagcggccctgcatac
gccctgagatctcaagaacagatgaccctttggccgcaaggggttgagctcttgacccagccagttcttgacccgagatttcctgcccgcgtctacgg
aaccagcagtttgtgcacttcacccagctgtcttacctgacctgtcttacctgacctgtcagccgagctgttatgaccgagaagtgaggagctgggggtgcggggtgcagtatacgggcaggga
tgctcccagctgcaggtgggagaaagtgaggaccaatgaccgagaaggagtgggggaggtgcggggtgcagtataacgggcaggga
cagcttcaaggctttcgccaaggctctgggtgtcatggatgaccttaagtcggggttccgagagctgctaccggggtattgtcacctt
ccagttcccgggccgccgtgtccacctgggcgccccaccgacgtgggaggggctatgatcctagctggaattag

FIG. 8

GnT1 Cys121Thr

avipilviacdrstvrrtldkilhyrpsaelfpiivsqdcgheetaqaiasygsavthirqpdlssiavppdhrkfqgyykiarhyrwal
gqvfrqfrfpaavvveddlevapdffeyfratypllkadpslwcvsawndngkeqmvdasrpellyrtdffpglgwlllaelwael
epkwpkafwddwmrpeqrqgracirpeisrtmtfgrkgvshgqffdqhlkfikinqqfvhftqldlsylqreaydrdflarvyg
apqlqvekvrtndrkelgevrvqytgrdsfkafakalgvmddlksgvpragyrgivtfqfpgrrvhlappptwegydpswn*

Gcggtgattccatcctggtcatcgcctgtgaccgcagcagctgttcggcgcagcagctgctgcattatcggccctcggctg
agctcttcccccatcatcttagccaggactgcgggggcacgcagagacggccccagccatcctcctacggcgcggtcacgcac
atccggcagccgacctgagcgagttcgggtgccgccgaccaccgcaagttccaggctactacaagatcgccgccactacc
gctgggcggctggggaaggtcttccggcagttcgcttcccgcggccgtggtgtggaggatgacctgagggtgcccggacttct
tcgagtactttcgggccacctatccgctgctgaaggccgaccctcccctgtggtgcgtctccgcctggctgtgttggccgagctctgggc
agatggtggacgcccaggcctgagctgctacccaccgcaccgcacctgggacgactggatgcgggccgccgagccgagcctgcatac
tgagctgagccccaagtggccaaaggccttctgggacgactggatgcgggccgccgagcagttctttgaccagcacctcaagtttatcaagctg
gccctgagatccaagaacgatgacctttgtgcacttcaccagctgacctgttcttacctgcagcggaggccatgaccgaagaggccataacggaagaggctgggggtgcagtatacgggcaggga
tgctcccccagctgcaggtggagaagtgaggaggaccaatgaccggaaggagctgggggtgcagtatacgggcaggga
cagcttcaaggctttcgccaaggctctggtgtcatgatgaccttaagtcgggggttccgagagcgtcagtatacgggcaggga
ccagttcccgggccgcctgtccacctggccgccccaccgacgtgggaggaggctatgatcctagctggaattag

FIG. 9

GnT1 Cys121Ala

avipilviacdrstvrraldkllhyrpsaelfpiivsqdcgheetaqaiasygsavthirqpdlssiavppdhrkfqgyykiarhyrwa
lgqvfrqfrfpaavvveddlevapdffeyfratypllkadpslwcvsawndngkeqmvdasrpellyrtdffpglgwlllaelwae
lepkwpkafwddwmrpeqrqgracirpeisrtmtfgrkgvshgqffdqhlkfiklnqqfvhftqldisylqreaydrdflarvyg
apqlqvekvrtndrkelgevrvqytgrdsfkafakalgvmddllksgypragyrgivtfqfpgrrvhlappptwegydpswn*

Gcggtgattccatcctggtcatcgcctgaccgcagcagcactgttcgcggcgccctagacaagctgctgcattatcggccctcggctg
agctcttcccatcatcgttagccaggactgcgggcacgaggagacgggcccaggccatcgcctcctacggcagcgcggtcacgcac
atccggcagcccgacctgagcagcattgcggtgccgcggccgccaccgccaagttccaggctactacaagatcgcgcgccactacc
gctgggcgctgggccaggtcttccggcagtttcgcttcccgccggccgtggtggtggatgacctggaggtggcccggacttct
tcgagtactttcgggccaccatccgctgctgaaggccgacccccctgtggtgcgtctcgcctggctggctgcgttggccgagtctgggc
agatggtggacgcgcagccaggcctgagcgctctaccgcaccgcaccgactgatgcggcggccgcgggcagcgcgggctggctgcatac
tgagctggagccccaagtggccaaaggccttcggcccgccaagggtgtgagcgcacgggcacgggcaggggctggctgcatac
gccctgagatctcaagaacgatgacctttggccgcaagaccgacacgggcggggaggagccgagatttcctcgcccggtctacgg
aacagcagtttgtgcacttcacccagctgacctggtggtgggtgcatgggagaaagtgaggaccaatgaccgggaaggagctgggga
tgctcccagctgcaggtggagaaagtgaggaccaatgaccgggaaggagctggggggtgggtggtcagtataacgggcaggga
cagcttcaaggctttcgccaaggctctggtgttcatggatgatgaccttaagtcgggtgtaccggggttccgagagctgctaccgggtattgtcacctt
ccagttcccgggccgcgccgtgtccaccctggccgccccaccgacgtgggaggggctatgatcctagctgggaattag

FIG. 10

GnT1 Arg120Ala, Cys121H

avipilviacdrstvrahldkilhyrpsaelfpiivsqdcgheetaqaiasygsavthirqpdlssiavppdhrkfqgyykiarhyrw
algqvfrqffpaavvveddlevapdffeyfratypllkadpslwcvsawndngkeqmvdasrpellyrtdffpglgwlllaelwa
elepkwpkafwddwmrpeqrqracirpeisrtmtfgrkgvshgqffdqhlkfiklnqqfvhftqldlsylqreaydrdflarvy
gapqlqvekvrtndrkelgevrvqytgrdsfkafakalgvmddlksgvpragyrgivtfqfpgrrvhlappptwegydpswn*

Gcggtgattccatcctggtcatcgcctgaccgcagcactgttcgggcccacctagacaagtctgctgcattatcggccctcggctg
agctcttcccatcatcgttagccaggactgcgggcacgagagacggccaggccatcgcctcctacggcagcgcggtcacgcac
atccggcagcccgacctgagcagcattgcggtgcgcgaccacccgcaagttccaggctactacaagatcgcgcgccactacc
gctgggcgctgggccagttctccggcagttgcttccccgcgccgtggtggaggatgacctgagggtggcccccgacttct
tcgagtacttcggggccacctatccgctgaaggccgaccctccctgttggtcgtctcggccctggggccttgggagctctgggc
agatggtggacgccagcaggcctgagctgctctaccgcaccgactttttccctgggacctgtctgttggccgagctctgggc
tgagctggagcccaagtggccaaaggccttctgctgggacgactgtggccgccaggtgtgagccacgggcggggcctgcatac
gccctgagatctcaagaacagatgaccttggccgcaaggttgtgagccacgggcagttcttgaccagccaagtttatcaagctg
aacagcagtttgtgcacttcacccagctggacctgtctacctgcagcggaggaggctgatggaccgagatttcctgccgcgctacgg
tgctcccagctgcaggtgaagaaagtgaggaccaatgaccggaaggagctgggggtgcggggtgcagtataacggcaggga
cagcttcaaggctttcgccaaggtctgcaaggtctgtcatgatgacctaagtcggggttccgagagctggtctaccgggtattgtcactt
ccagttcccgcggccgtccacctggcgccccaccgacgtgggaggctatgatcctagctggaattag

FIG. 11

Rat Liver ST3Gal III amino acid sequence:

MGLLVFVRNLLLALCLFLVLGFLYTSAWKLHLLQWEDSNSLILSLDSAGQTLGTEYDRL
GFLLKLDSKLPAELATKYANFSEGACKPGYASAMMTAIFPRFSKPAPMFLDDSFRKW
ARIREFVPPFGIKGQDNLIKAILSVTKEYRLTPALDSLHCRRCIIVGNGGVLANKSLGS
RIDDYDIVIRLNSAPVKGFEKDVGSKTTLRITYPEGAMQRPEQYERDSLFVLAGFKW
QDFKWLKYTVYKERVSASDGFWKSVATRVPKEPPERILNPYFIQEAAFTLIGLPFNN
GLMGRGNIPTLGSVAVTMALDGCDEVAVAGFGYDMNTPNAPLHYYETVRMAAIKE
SWTHNIQREKEFLRKLVKARVITDLSSGI

FIG. 12

Full length UDP-N-acetylgalactosaminyltransferase 2 (GalNAcT2) nucleic acid and amino acid sequences

Amino acid sequence

```
Met Arg Arg Arg Ser Arg Met Leu Leu Cys Phe Ala Phe Leu Trp Val
1           5               10              15

Leu Gly Ile Ala Tyr Tyr Met Tyr Ser Gly Gly Gly Ser Ala Leu Ala
        20              25              30

Gly Gly Ala Gly Gly Gly Ala Gly Arg Lys Glu Asp Trp Asn Glu Ile
        35              40              45

Asp Pro Ile Lys Lys Lys Asp Leu His His Ser Asn Gly Glu Glu Lys
    50              55              60

Ala Gln Ser Met Glu Thr Leu Pro Pro Gly Lys Val Arg Trp Pro Asp
65              70              75              80

Phe Asn Gln Glu Ala Tyr Val Gly Gly Thr Met Val Arg Ser Gly Gln
            85              90              95

Asp Pro Tyr Ala Arg Asn Lys Phe Asn Gln Val Glu Ser Asp Lys Leu
        100             105             110

Arg Met Asp Arg Ala Ile Pro Asp Thr Arg His Asp Gln Cys Gln Arg
        115             120             125

Lys Gln Trp Arg Val Asp Leu Pro Ala Thr Ser Val Val Ile Thr Phe
    130             135             140

His Asn Glu Ala Arg Ser Ala Leu Leu Arg Thr Val Val Ser Val Leu
145             150             155             160

Lys Lys Ser Pro Pro His Leu Ile Lys Glu Ile Ile Leu Val Asp Asp
            165             170             175

Tyr Ser Asn Asp Pro Glu Asp Gly Ala Leu Leu Gly Lys Ile Glu Lys
        180             185             190

Val Arg Val Leu Arg Asn Asp Arg Arg Glu Gly Leu Met Arg Ser Arg
        195             200             205

Val Arg Gly Ala Asp Ala Ala Gln Ala Lys Val Leu Thr Phe Leu Asp
    210             215             220

Ser His Cys Glu Cys Asn Glu His Trp Leu Glu Pro Leu Leu Glu Arg
225             230             235             240

Val Ala Glu Asp Arg Thr Arg Val Val Ser Pro Ile Ile Asp Val Ile
            245             250             255

Asn Met Asp Asn Phe Gln Tyr Val Gly Ala Ser Ala Asp Leu Lys Gly
        260             265             270

Gly Phe Asp Trp Asn Leu Val Phe Lys Trp Asp Tyr Met Thr Pro Glu
    275             280             285
```

FIG. 13A

14/54

Gln Arg Arg Ser Arg Gln Gly Asn Pro Val Ala Pro Ile Lys Thr Pro
   290   295   300

Met Ile Ala Gly Gly Leu Phe Val Met Asp Lys Phe Tyr Phe Glu Glu
305   310   315   320

Leu Gly Lys Tyr Asp Met Met Met Asp Val Trp Gly Gly Glu Asn Leu
            325   330   335

Glu Ile Ser Phe Arg Val Trp Gln Cys Gly Gly Ser Leu Glu Ile Ile
        340   345   350

Pro Cys Ser Arg Val Gly His Val Phe Arg Lys Gln His Pro Tyr Thr
        355   360   365

Phe Pro Gly Gly Ser Gly Thr Val Phe Ala Arg Asn Thr Arg Arg Ala
   370   375   380

Ala Glu Val Trp Met Asp Glu Tyr Lys Asn Phe Tyr Tyr Ala Ala Val
385   390   395   400

Pro Ser Ala Arg Asn Val Pro Tyr Gly Asn Ile Gln Ser Arg Leu Glu
          405   410   415

Leu Arg Lys Lys Leu Ser Cys Lys Pro Phe Lys Trp Tyr Leu Glu Asn
        420   425   430

Val Tyr Pro Glu Leu Arg Val Pro Asp His Gln Asp Ile Ala Phe Gly
        435   440   445

Ala Leu Gln Gln Gly Thr Asn Cys Leu Asp Thr Leu Gly His Phe Ala
   450   455   460

Asp Gly Val Val Gly Val Tyr Glu Cys His Asn Ala Gly Gly Asn Gln
465   470   475   480

Glu Trp Ala Leu Thr Lys Glu Lys Ser Val Lys His Met Asp Leu Cys
            485   490   495

Leu Thr Val Val Asp Arg Ala Pro Gly Ser Leu Ile Lys Leu Gln Gly
        500   505   510

Cys Arg Glu Asn Asp Ser Arg Gln Lys Trp Glu Gln Ile Glu Gly Asn
        515   520   525

Ser Lys Leu Arg His Val Gly Ser Asn Leu Cys Leu Asp Ser Arg Thr
   530   535   540

Ala Lys Ser Gly Gly Leu Ser Val Glu Val Cys Gly Pro Ala Leu Ser
545   550   555   560

Gln Gln Trp Lys Phe Thr Leu Asn Leu Gln Gln
        565   570

FIG. 13A (CONT.)
SUBSTITUTE SHEET (RULE 26)

Nucleic acid sequence

```
atgggggggc gctgcgggat gctgctctgc ttcgccttcc tgtgggtgct gggcatcgcc      60

tactacatgt actcgggggg cggctctgcg ctggccgggg gcgcgggcgg cggcgccggc     120

aggaaggagg actggaatga aattgacccc attaaaaaga aagacttca  tcacagcaat     180

ggagaagaga aagcacaaag catggagacc ctccctccag ggaaagtacg gtggccagac     240

tttaaccagg aagcttatgt tggagggacg atggtccgct ccgggcagga cccttacgcc     300

cgcaacaagt tcaaccaggt ggagagtgat aagcttcgaa tggacagagc catccctgac     360

acccggcatg accagtgtca gcggaagcag tggcgggtgg atctgccggc caccagcgtg     420

gtgatcacgt ttcacaatga agccaggtcg gccctactca ggaccgtggt cagcgtgctt     480

aagaaaagcc cgccccatct cataaaagaa atcatcttgg tggatgacta cagcaatgat     540

cctgaggacg gggctctctt ggggaaaatt gagaaagtgc gagttcttag aaatgatcga     600

cgagaaggcc tcatgcgctc acgggttcgg ggggccgatg ctgcccaagc caaggtcctg     660

aacttcctgg acagtcactg cgagtgtaat gagcactggc tggagcccct cctggaaagg     720

gtggcggagg acaggactcg ggttgtgtca cccatcatcg atgtcattaa tatggacaac     780

tttcagtatg tggggggcatc tgctgacttg aagggcggtt ttgattggaa cttggtattc     840

aagtgggatt acatgacgcc tgagcagaga aggtcccggc aggggaaccc agtcgcccct     900

ataaaaaccc ccatgattgc tggtgggctg tttgtgatgg ataagttcta ttttgaagaa     960

ctggggaagt acgacatgat gatggatgtg tggggaggag agaacctaga gatcctgttc    1020

cgcgtgtggc agtgtggtgg cagcctggag atcatcccgt gcagccgtgt gggacacgtg    1080

ttccggaagc agcacccta  cacgttcccg ggtggcagtg gcactgtctt tgcccgaaac    1140

acccgccggg cagcagaggt ctggatggat gaatacaaaa atttctatta tgcagcagtg    1200

ccttctgcta gaaacgttcc ttatggaaat attcagagca gattggagct taggaagaaa    1260

ctcagctgca agcctttcaa atggtacctt gaaaatgtct atccagagtt aagggttcca    1320

gaccatcagg atatagcttt tggggccttg cagcagggaa ctaactgcct cgacactttg    1380

ggacactttg ctgatggtgt ggttggagtt tatgaatgtc acaatgctgg gggaaaccag    1440

gaatgggcct tgacgaagga gaagtcggtg aagcacatgg atttgtgcct tactgtggtg    1500

gaccgggcac cgggctctct tataaagctg cagggctgcc gagaaaatga cagcagacag    1560

aaatgggaac agatcgaggg caactccaag ctgaggcacg tgggcagcaa cctgtgcctg    1620

gacagtcgca cggccaagag cggggggccta agcgtggagg tgtgtggccc ggccttttcg    1680

cagcagtgga agttcacgct caacctgcag cag                                  1713
```

FIG. 13B

16/54

Δ51 UDP-N-acetylgalactosaminyltransferase 2, GalNAcT2, nucleic acid and amino acid sequences

Amino acid sequence

Lys Lys Lys Asp Leu His His Ser Asn Gly Glu Glu Lys Ala Gln Ser
1   5   10   15

Met Glu Thr Leu Pro Pro Gly Lys Val Arg Trp Pro Asp Phe Asn Gln
              20   25   30

Glu Ala Tyr Val Gly Gly Thr Met Val Arg Ser Gly Gln Asp Pro Tyr
         35   40   45

Ala Arg Asn Lys Phe Asn Gln Val Glu Ser Asp Lys Leu Arg Met Asp
    50   55   60

Arg Ala Ile Pro Asp Thr Arg His Asp Gln Cys Gln Arg Lys Gln Trp
65   70   75   80

Arg Val Asp Leu Pro Ala Thr Ser Val Val Ile Thr Phe His Asn Glu
         85   90   95

Ala Arg Ser Ala Leu Leu Arg Thr Val Val Ser Val Leu Lys Lys Ser
         100   105   110

Pro Pro His Leu Ile Lys Glu Ile Ile Leu Val Asp Asp Tyr Ser Asn
         115   120   125

Asp Pro Glu Asp Gly Ala Leu Leu Gly Lys Ile Glu Lys Val Arg Val
    130   135   140

Leu Arg Asn Asp Arg Arg Glu Gly Leu Met Arg Ser Arg Val Arg Gly
145   150   155   160

Ala Asp Ala Ala Gln Ala Lys Val Leu Thr Phe Leu Asp Ser His Cys
              165   170   175

Glu Cys Asn Glu His Trp Leu Glu Pro Leu Leu Glu Arg Val Ala Glu
         180   185   190

Asp Arg Thr Arg Val Val Ser Pro Ile Ile Asp Val Ile Asn Met Asp
         195   200   205

Asn Phe Gln Tyr Val Gly Ala Ser Ala Asp Leu Lys Gly Gly Phe Asp
    210   215   220

Trp Asn Leu Val Phe Lys Trp Asp Tyr Met Thr Pro Glu Gln Arg Arg
225   230   235   240

Ser Arg Gln Gly Asn Pro Val Ala Pro Ile Lys Thr Pro Met Ile Ala
              245   250   255

Gly Gly Leu Phe Val Met Asp Lys Phe Tyr Phe Glu Glu Leu Gly Lys
         260   265   270

Tyr Asp Met Met Met Asp Val Trp Gly Gly Glu Asn Leu Glu Ile Ser
    275   280   285

Phe Arg Val Trp Gln Cys Gly Gly Ser Leu Glu Ile Ile Pro Cys Ser

FIG. 14A
SUBSTITUTE SHEET (RULE 26)

```
        290                    295                    300

Arg Val Gly His Val Phe Arg Lys Gln His Pro Tyr Thr Phe Pro Gly
305             310             315             320

Gly Ser Gly Thr Val Phe Ala Arg Asn Thr Arg Arg Ala Ala Glu Val
            325             330             335

Trp Met Asp Glu Tyr Lys Asn Phe Tyr Tyr Ala Ala Val Pro Ser Ala
            340             345             350

Arg Asn Val Pro Tyr Gly Asn Ile Gln Ser Arg Leu Glu Leu Arg Lys
            355             360             365

Lys Leu Ser Cys Lys Pro Phe Lys Trp Tyr Leu Glu Asn Val Tyr Pro
370             375             380

Glu Leu Arg Val Pro Asp His Gln Asp Ile Ala Phe Gly Ala Leu Gln
385             390             395             400

Gln Gly Thr Asn Cys Leu Asp Thr Leu Gly His Phe Ala Asp Gly Val
            405             410             415

Val Gly Val Tyr Glu Cys His Asn Ala Gly Gly Asn Gln Glu Trp Ala
            420             425             430

Leu Thr Lys Glu Lys Ser Val Lys His Met Asp Leu Cys Leu Thr Val
            435             440             445

Val Asp Arg Ala Pro Gly Ser Leu Ile Lys Leu Gln Gly Cys Arg Glu
    450             455             460

Asn Asp Ser Arg Gln Lys Trp Glu Gln Ile Glu Gly Asn Ser Lys Leu
465             470             475             480

Arg His Val Gly Ser Asn Leu Cys Leu Asp Ser Arg Thr Ala Lys Ser
            485             490             495

Gly Gly Leu Ser Val Glu Val Cys Gly Pro Ala Leu Ser Gln Gln Trp
            500             505             510

Lys Phe Thr Leu Asn Leu Gln Gln
            515             520
```

FIG. 14A (CONT.)

Nucleic acid sequence   18/54

aaaaagaaag	aacttcatca	cagcaatgga	gaagagaaag	cacaaagcat	ggagaacctc	60
cctccaggga	aagtacggtg	gccagacttt	aaccaggaag	cttatgttgg	agggacgatg	120
gtccgctccg	ggcaggaccc	ttacgcccgc	aacaagttca	accaggtgga	gagtgataag	180
cttcgaatgg	acagagccat	ccctgacacc	cggcatgacc	agtgtcagcg	gaagcagtgg	240
cgggtggatc	tgccggccac	cagcgtggtg	atcacgtttc	acaatgaagc	caggtcggcc	300
ctactcagga	ccgtggtcag	cgtgcttaag	aaaagcccgc	cccatctcat	aaaagaaatc	360
atcttggtgg	atgactacag	caatgatcct	gaggacgggg	ctctcttggg	gaaaattgag	420
aaagtgcgag	ttcttagaaa	tgatcgacga	gaaggcctca	tgcgctcacg	ggttcggggg	480
gccgatgctg	cccaagccaa	ggtcctgacc	ttcctggaca	gtcactgcga	gtgtaatgag	540
cactggctgg	agcccctcct	ggaaagggtg	gcggaggaca	ggactcgggt	tgtgtcaccc	600
atcatcgatg	tcattaatat	ggacaacttt	cagtatgtgg	gggcatctgc	tgacttgaag	660
ggcggttttg	attggaactt	ggtattcaag	tgggattaca	tgacgcctga	gcagagaagg	720
tcccggcagg	ggaacccagt	cgcccctata	aaaaccccca	tgattgctgg	tgggctgttt	780
gtgatggata	agttctattt	tgaagaactg	gggaagtacg	acatgatgat	ggatgtgtgg	840
ggaggagaga	acctagagat	ctcgttccgc	gtgtggcagt	gtggtggcag	cctggagatc	900
atcccgtgca	gccgtgtggg	acacgtgttc	cggaagcagc	acccctacac	gttccggggt	960
ggcagtggca	ctgtctttgc	ccgaaacacc	cgccgggcag	cagaggtctg	gatggatgaa	1020
tacaaaaatt	tctattatgc	agcagtgcct	tctgctagaa	acgttcctta	tggaaatatt	1080
cagagcagat	tggagcttag	gaagaaactc	agctgcaagc	ctttcaaatg	gtaccttgaa	1140
aatgtctatc	cagagttaag	ggttccagac	catcaggata	tagcttttgg	ggccttgcag	1200
cagggaacta	actgcctcga	cactttggga	cactttgctg	atggtgtggt	tggagtttat	1260
gaatgtcaca	atgctggggg	aaaccaggaa	tgggcttga	cgaaggagaa	gtcggtgaag	1320
cacatggatt	tgtgccttac	tgtggtggac	cgggcaccgg	gctctcttat	aaagctgcag	1380
ggctgccgag	aaaatgacag	cagacagaaa	tgggaacaga	tcgagggcaa	ctccaagctg	1440
aggcacgtgg	gcagcaacct	gtgcctggac	agtcgcacgg	ccaagagcgg	gggcctaagc	1500
gtggaggtgt	gtggcccggc	cctttcgcag	cagtggaagt	tcacgctcaa	cctgcagcag	1560

FIG. 14B

Protein Concentration of refolded MBP-GalNacT2(Δ51)
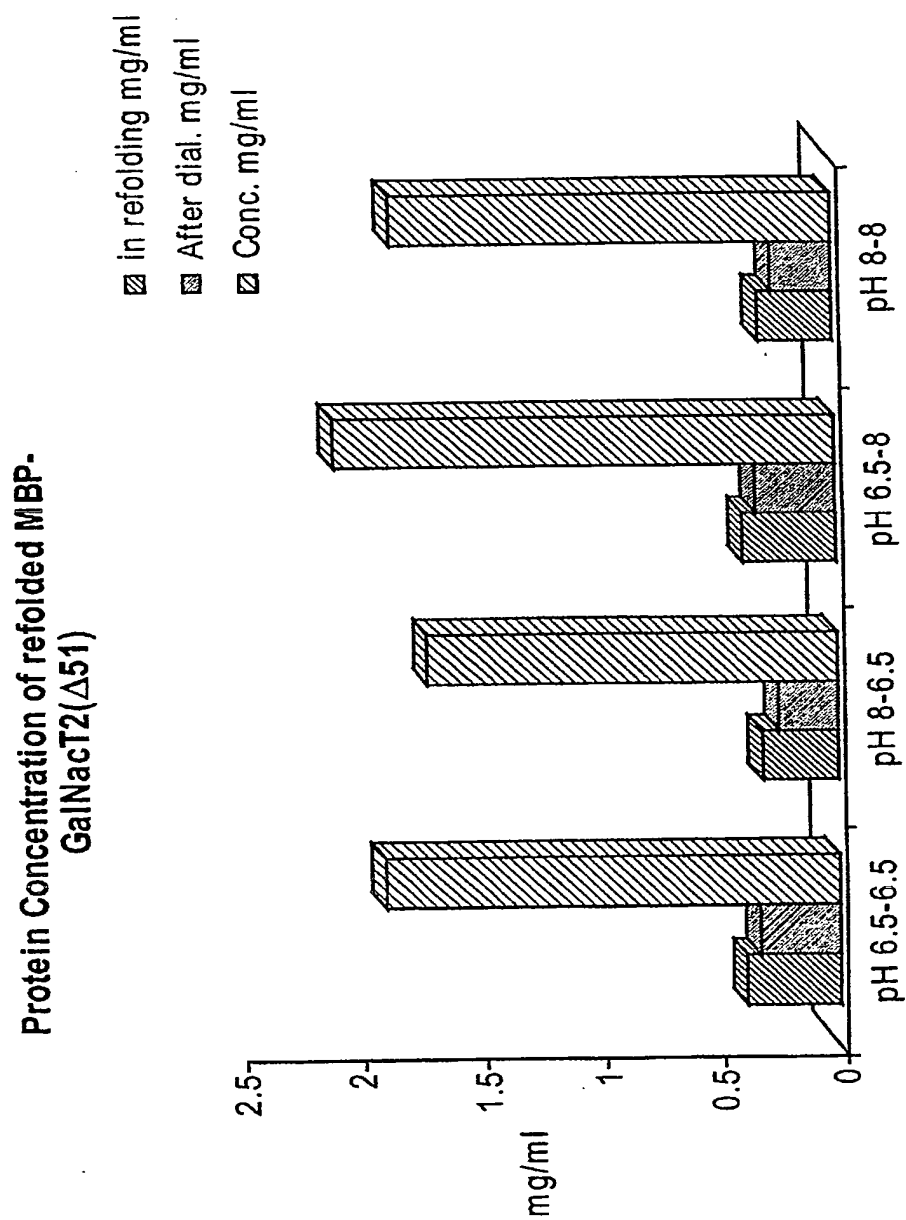
FIG. 15

GaINAcT2 activities of refolded MBP-GaINAcT2(D51)

FIG. 16

pH effect on the MBP-GalNAcT2(D51) specific activities
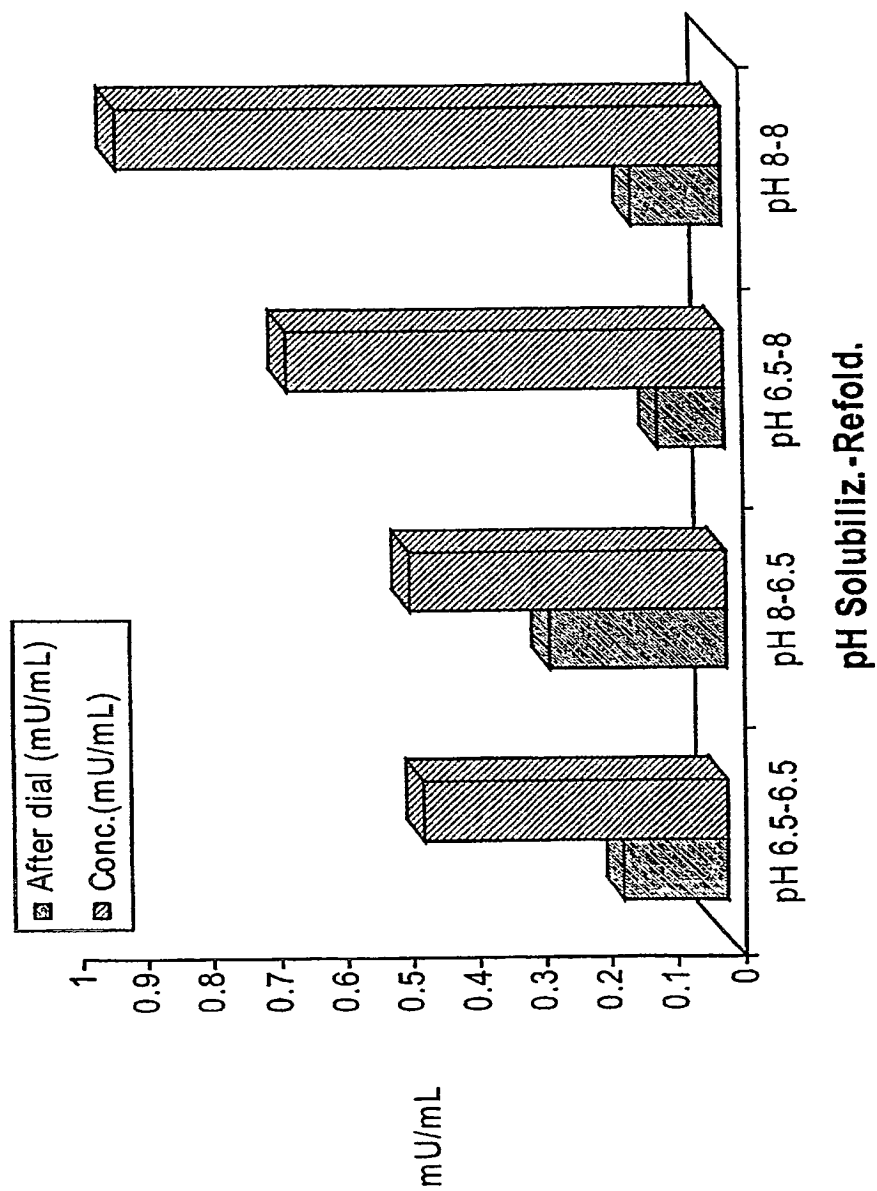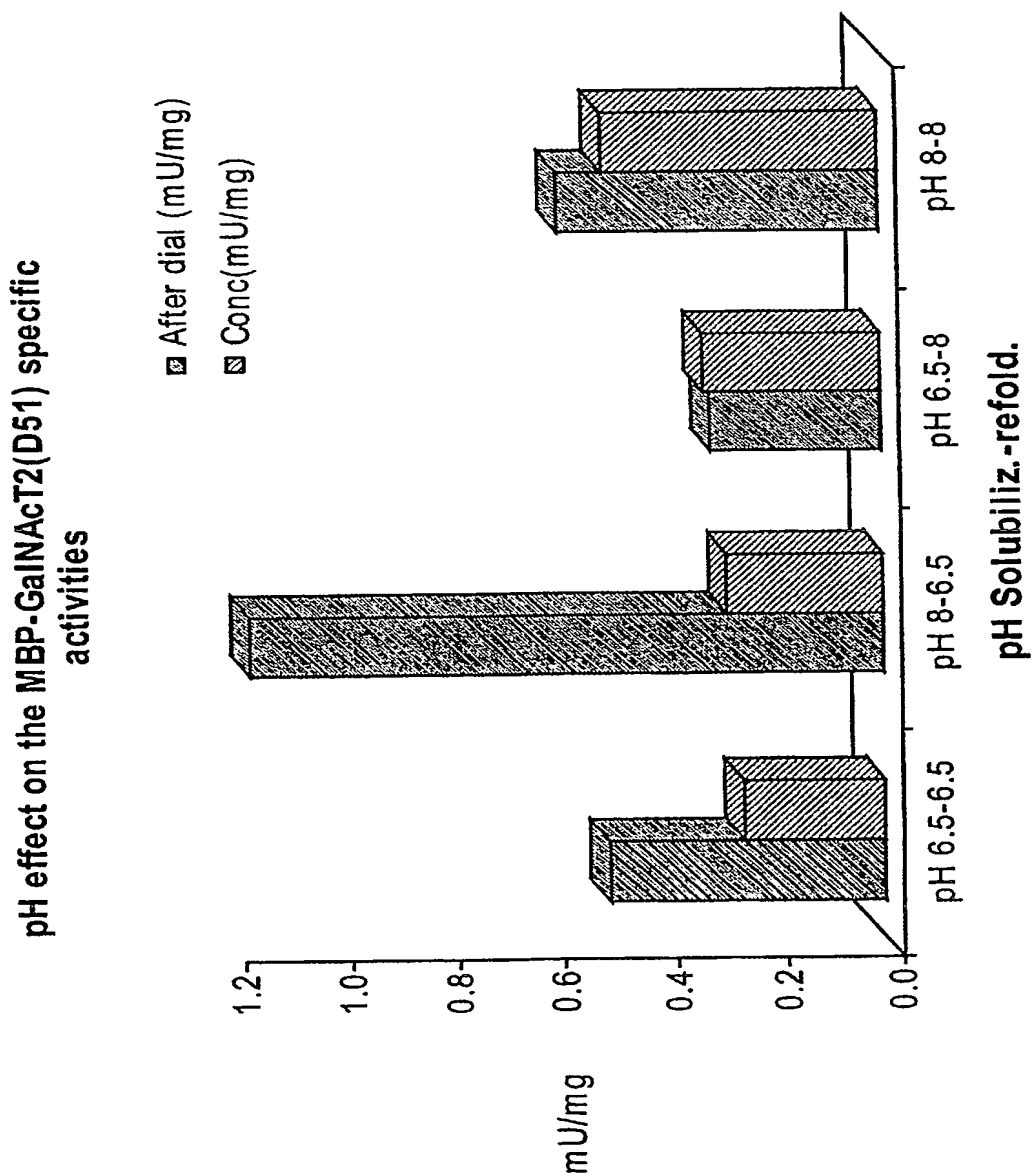
FIG. 17
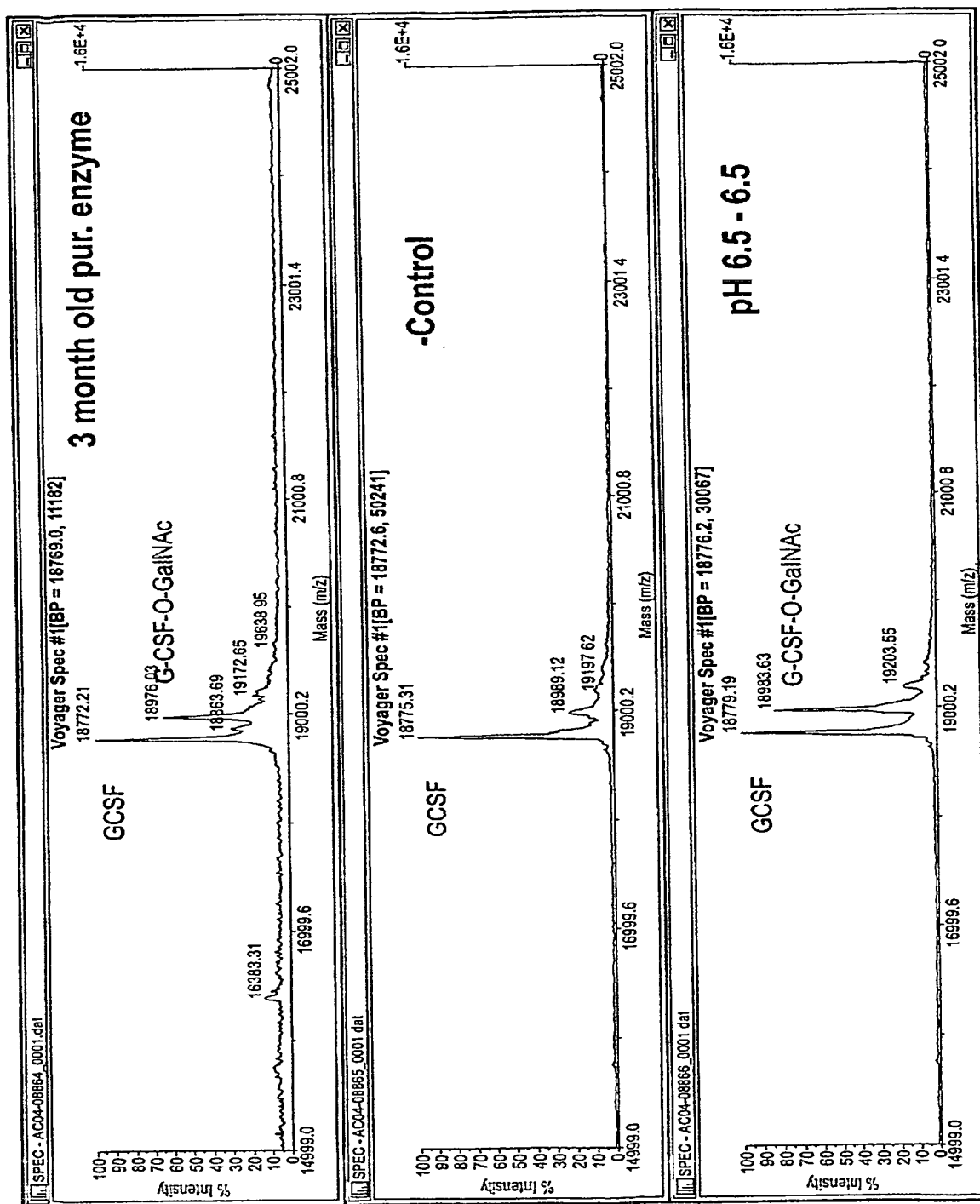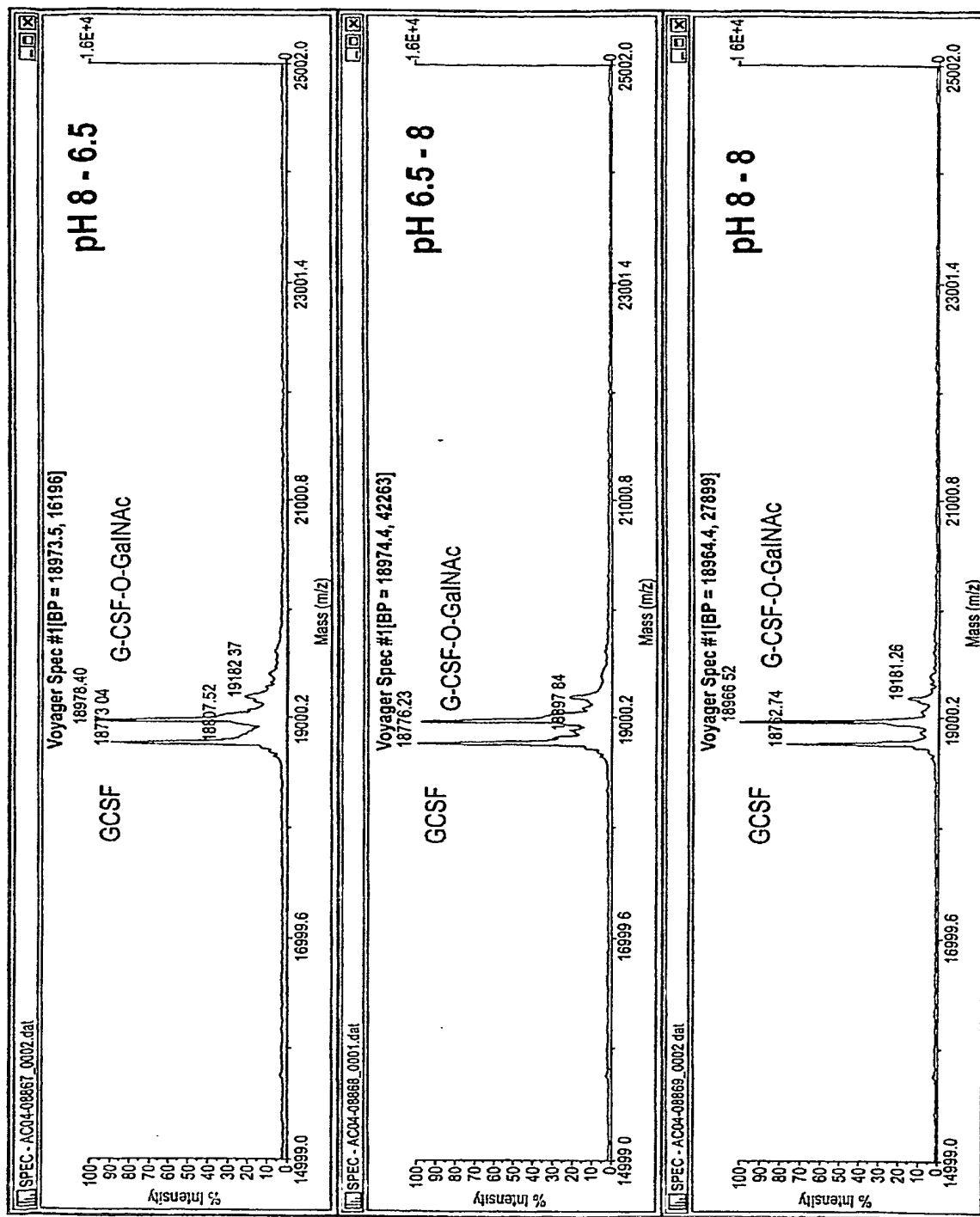
FIG. 18A

FIG. 18B
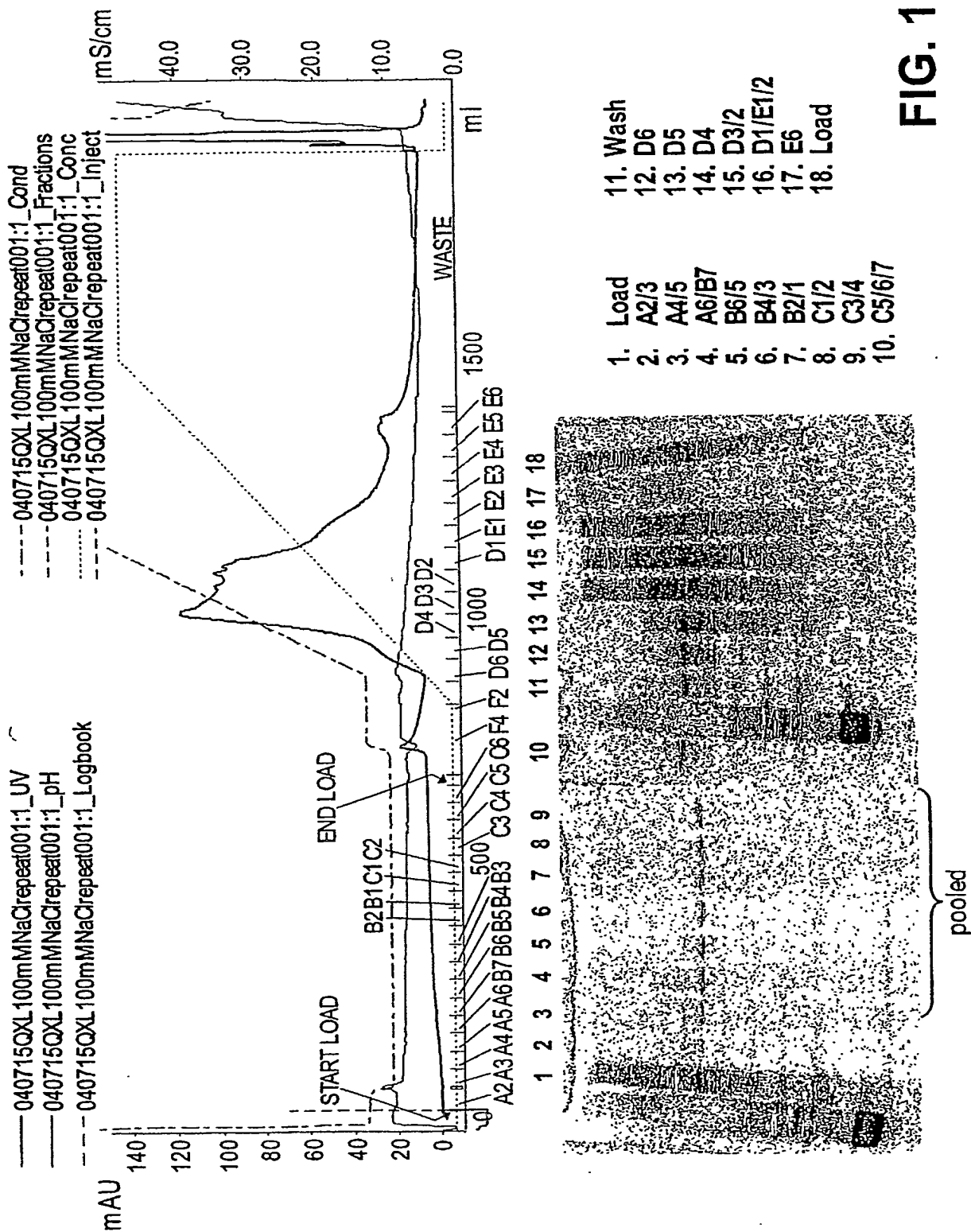
FIG. 19

	Volume (mL)	Activity (U/L)	A280
Load	890	1.5	0.110
FT$_{A4-C4}$	670	9.2	NA
FT$_{C5-C7}$	120	1.0	0
Wash	138	3.6	0.100
D6	45	4.5	0
D5	45	2.4	0.026
D4	45	2.0	0.108
D3/2	90	1.1	0.179
E6		0.0	0.017

FIG. 20
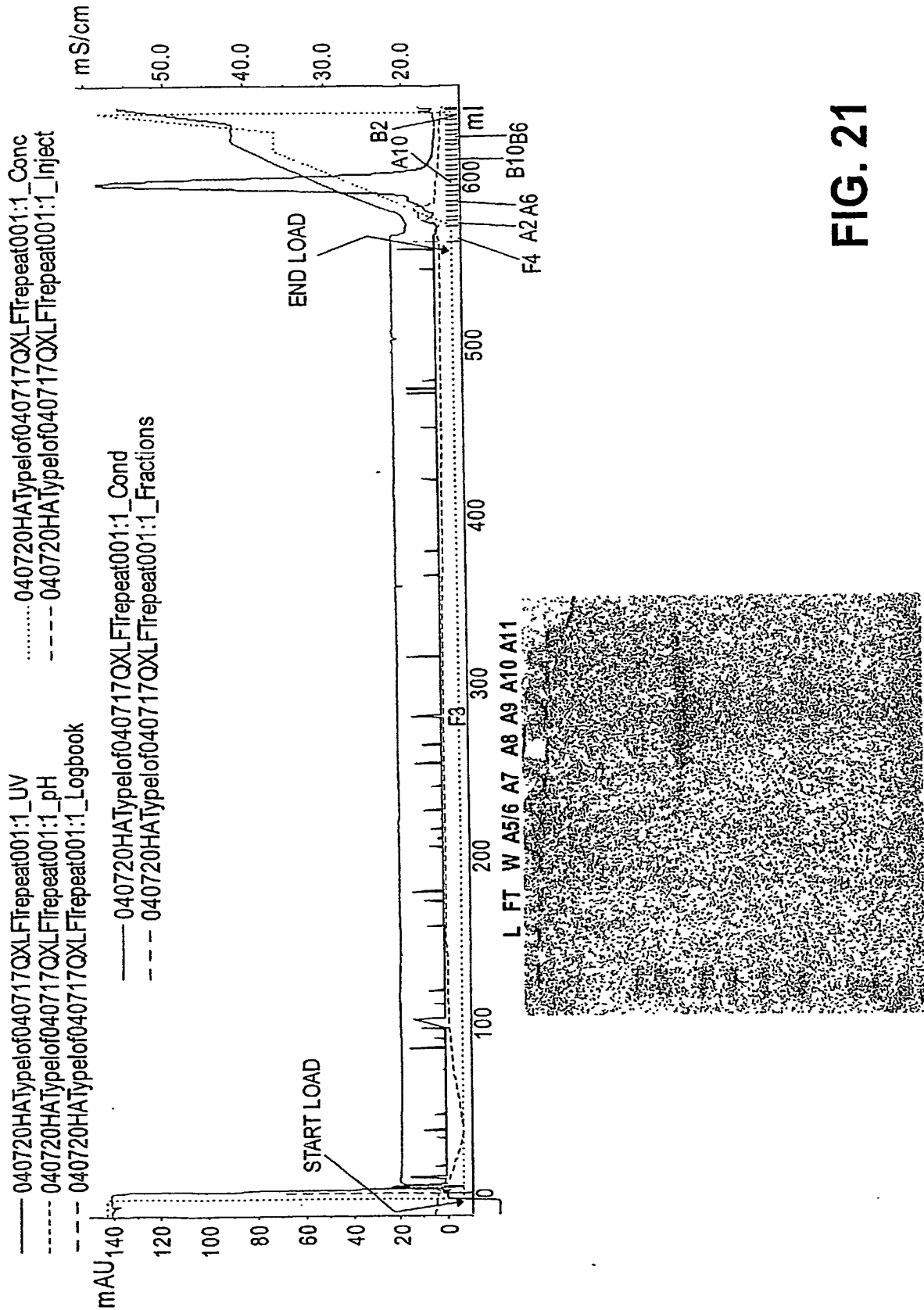
FIG. 21

Pooled A4-C4 and adjusted pH to 7.0 using 1mM HCl 670mL
load pH 7.0 16mS/cm

	Volume (mL)	Activity (U/L) pre-dialysis	Activity (U/L) post-dialysis	A280	A280/ 1.51 (mg/mL)	Activity (U)	Mass (mg)	Specific Activity (U/mg)
Load	670	9.2	NA	NA				
FT	670	0.0	NA	0.122				
Wash	9	2.9	NA	-0.013				
A5/6	6	1.1	NA	-0.005				
A7	3	0.1						
A8	3	1.3	19.3 (13mL)	0.180	0.119	0.25	1.55	0.16
A9	3	4.6						
A10	3	2.4						
A11	3	0.4						

FIG. 22

FIG. 23

MBP-pST3Gal1 fusion protein

MKIEEGKLVIWINGDKGYNGLAEVGKKFEKDTGIKVTVEHPDKLEEKFPQVAATGDGPDIIFWAHDRF
GGYAQSGLLAEITPDKAFQDKLYPFTWDAVRYNGKLIAYPIAVEALSLIYNKDLLPNPPKTWEEIPALD
KELKAKGKSALMFNLQEPYFTWPLIAADGGYAFKYENGKYDIKDVGVDNAGAKAGLTFLVDLIKNKH
MNADTDYSIAEAAFNKGETAMTINGPWAWSNIDTSKVNYGVTVLPTFKGQPSKPFVGVLSAGINAASP
NKELAKEFLENYLLTDEGLEAVNKDKPLGAVALKSYEEELAKDPRIAATMENAQKGEIMPNIPQMSAF
WYAVRTAVINAASGRQTVDEALKDAQTNSSSNNNNNNNNLGIEGRISEFGSELSENFKKLMKYPYR
PCTCTRCIEEQRVSAWFDERFNRSMQPLLTAKNAHLEEDTYKWWLRLQREKQPNNLNDTIRELFQVVP
GNVDPLLEKRLVSCRRCAVVGNSGNLKESYYGPQIDSHDFVLRMNKAPTEGFEADVGSKTTHHFVYPE
SFRELAQEVSMILVPFKITDLEWVISATTIGRISHTYVPVPAKIKVKKEKILIYHPAFIKYVFDRWLQGH
GRYPSTGILSVIFSLHICDEVDLYGFGADSKGNWHHYWENNPSAGAFRKTGVHDGDFESNVTTILASIN
KIRIFKGR

FIG. 24A
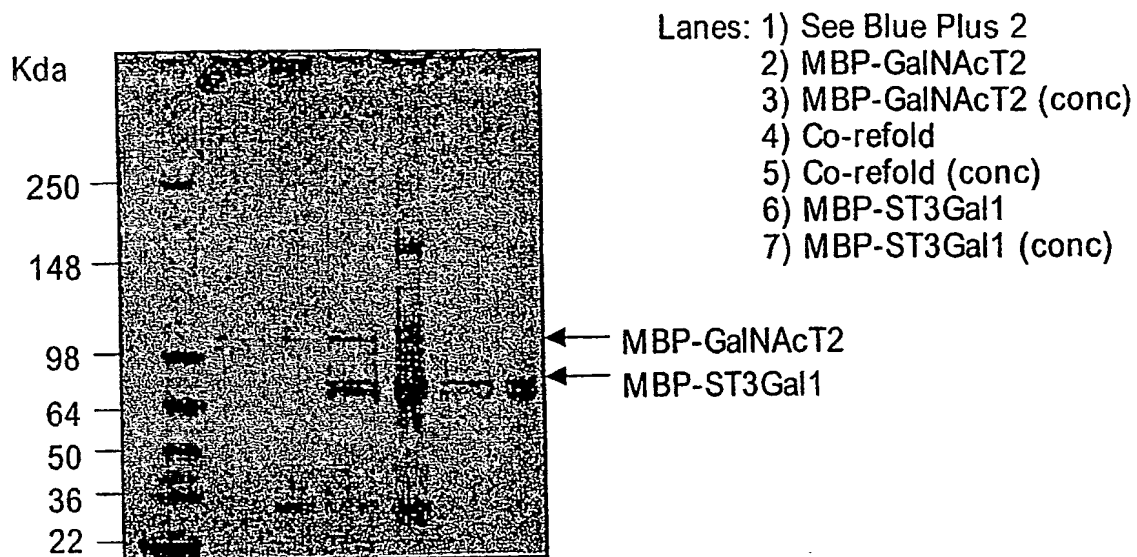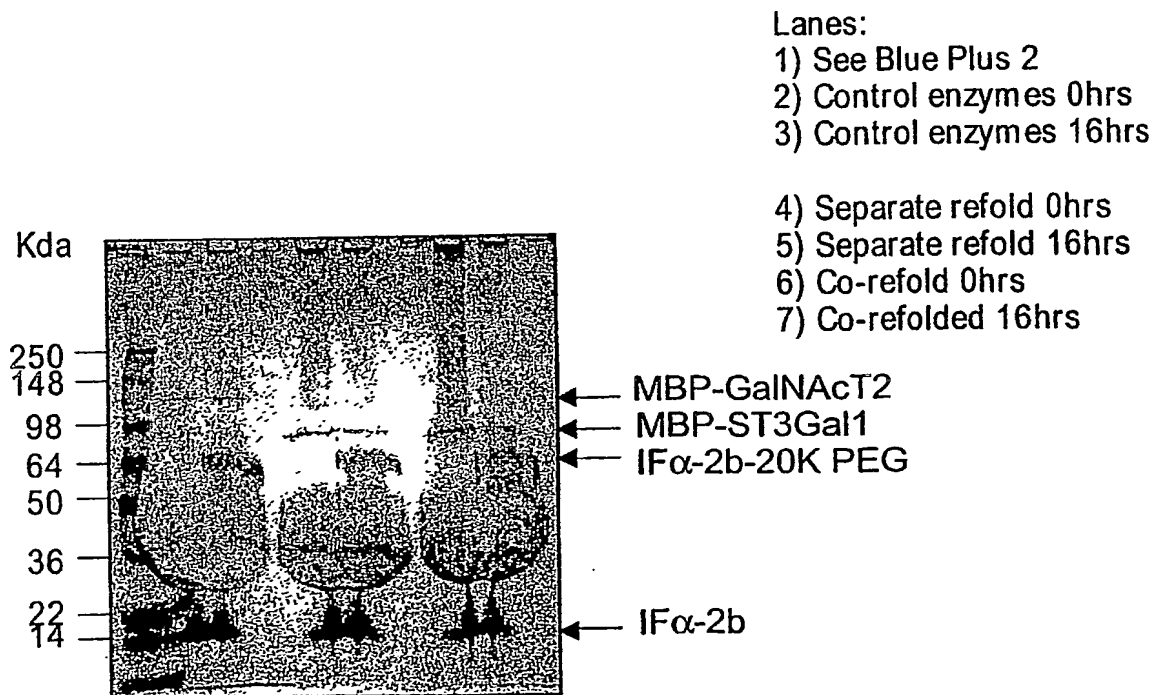
MBPSBD-pST3Gal1 fusion protein

MKIEEGKLVIWINGDKGYNGLAEVGKKFEKDTGIKVTVEHPDKLEEKFPQVAATGDGPDIIFWAHDRF
GGYAQSGLLAEITPDKAFQDKLYPFTWDAVRYNGKLIAYPIAVEALSLIYNKDLLPNPPKTWEEIPALD
KELKAKGKSALMFNLQEPYFTWPLIAADGGYAFKYENGKYDIKDVGVDNAGAKAGLTFLVDLIKNKH
MNADTDYSIAEAAFNKGETAMTINGPWAWSNIDTSKVNYGVTVLPTFKGQPSKPFVGVLSAGINAASP
NKELAKEFLENYLLTDEGLEAVNKDKPLGAVALKSYEEELAKDPRIAATMENAQKGEIMPNIPQMSAF
WYAVRTAVINAASGRQTVDEALKDAQTNSSSNNNNNNNNLGIEGRISEFGSIVAIGGTTTAIPTG
SGSVTSTSKTTATASKTSTSTSSCTTPTAVAVTFDLTATTYGENIYLVGSISQLGDWETSDGIALSAD
KYTSSDPLWYTVTLPAGESFEYKFIRIESDDSVEWESDPNREYTVPQACGTSTATVDTWRGSELSEN
FKKLMKYPYRPCTCTRCIEEQRVSAWFDERFNRSMQPLLTAKNAHLEEDTYKWWLRLQREKQPNNLN
DTIRELFQVVPGNVDPLLEKRLVSCRRCAVVGNSGNLKESYYGPQIDSHDFVLRMNKAPTEGFEADVG
SKTTHHFVYPESFRELAQEVSMILVPFKITDLEWVISATTIGRISHTYVPVPAKIKVKKEKILIYHPAFIK
YVFDRWLQGHGRYPSTGILSVIFSLHICDEVDLYGFGADSKGNWHHYWENNPSAGAFRKTGVHDGDF
ESNVTTILASINKIRIFKGR

FIG. 24B

Sialyltransferase Activity
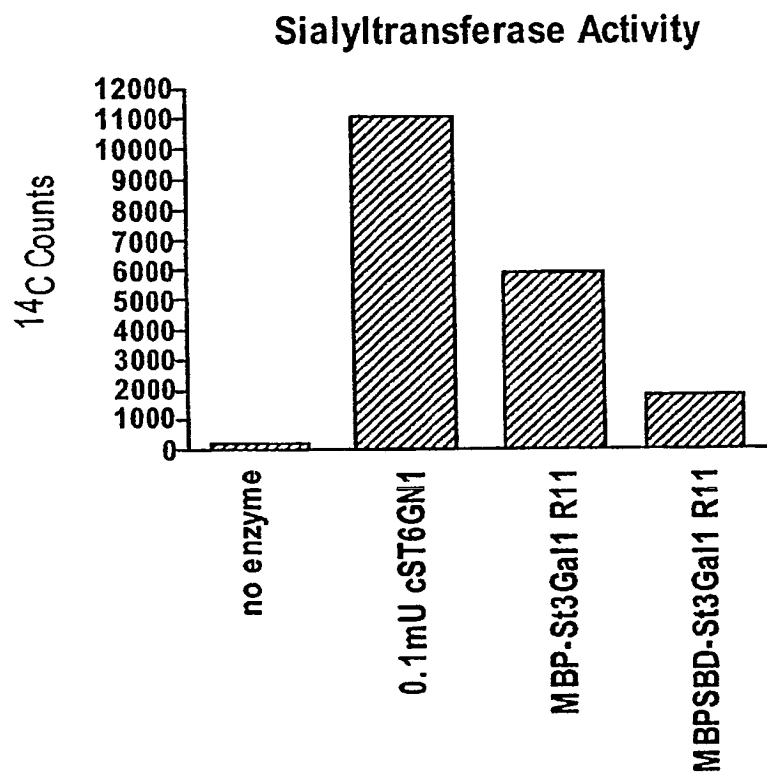
Test pST3Gal1 for activity after Hampton refold

FIG. 25

FIG. 26A

MKIEEGKLVIWINGDKGYNGLAEVGKKFEKDTGIKVTVEHPDKLEEKFPQVAATGDGPDIIFWAHDRF
GGYAQSGLLAEITPDKAFQDKLYPFTWDAVRYNGKLIAYPIAVEALSLIYNKDLLPNPPKTWEEIPALD
KELKAKGKSALMFNLQEPYFTWPLIAADGGYAFKYENGKYDIKDVGVDNAGAKAGLTFLVDLIKNKH
MNADTDYSIAEAAFNKGETAMTINGPWAWSNIDTSKVNYGVTVLPTFKGQPSKPFVGVLSAGINAASP
NKELAKEFLENYLLTDEGLEAVNKDKPLGAVALKSYEEELAKDPRIAATMENAQKGEIMPNIPQMSAF
WYAVRTAVINAASGRQTVDEALKDAQTNSSSNNNNNNNLGEGRISEFGSSEHLDKVPRITPGAL
STRKTPMATGAVPAKKKVVQATKSPASSPHPTTRRQRLKASEFKSEPRWDFEEEYSLDMSSLQT
NCSASVKIKASKSPWLQNIFLPNITLFLDSGRFTQSEWNRLEHFAPPFGFMELNQSLVQKVVTRFP
PVRQQQLLASLPTGYSKCITCAVVGNGGILNDSRVGREIDSHDYFERLSGAVIKGYEQDVGTRT
SFYGFTAFSLTQSILILGRRGFQHVPLGKDVRYLHFLEGTRNYEWLEAMFLNQTLAKTHLSWFR
HRPQEAFRNALDLDRYLLLHPDFLRYMKNRFLRSKTLDTAXWRIYRPTTGALLLTALHLCDKV
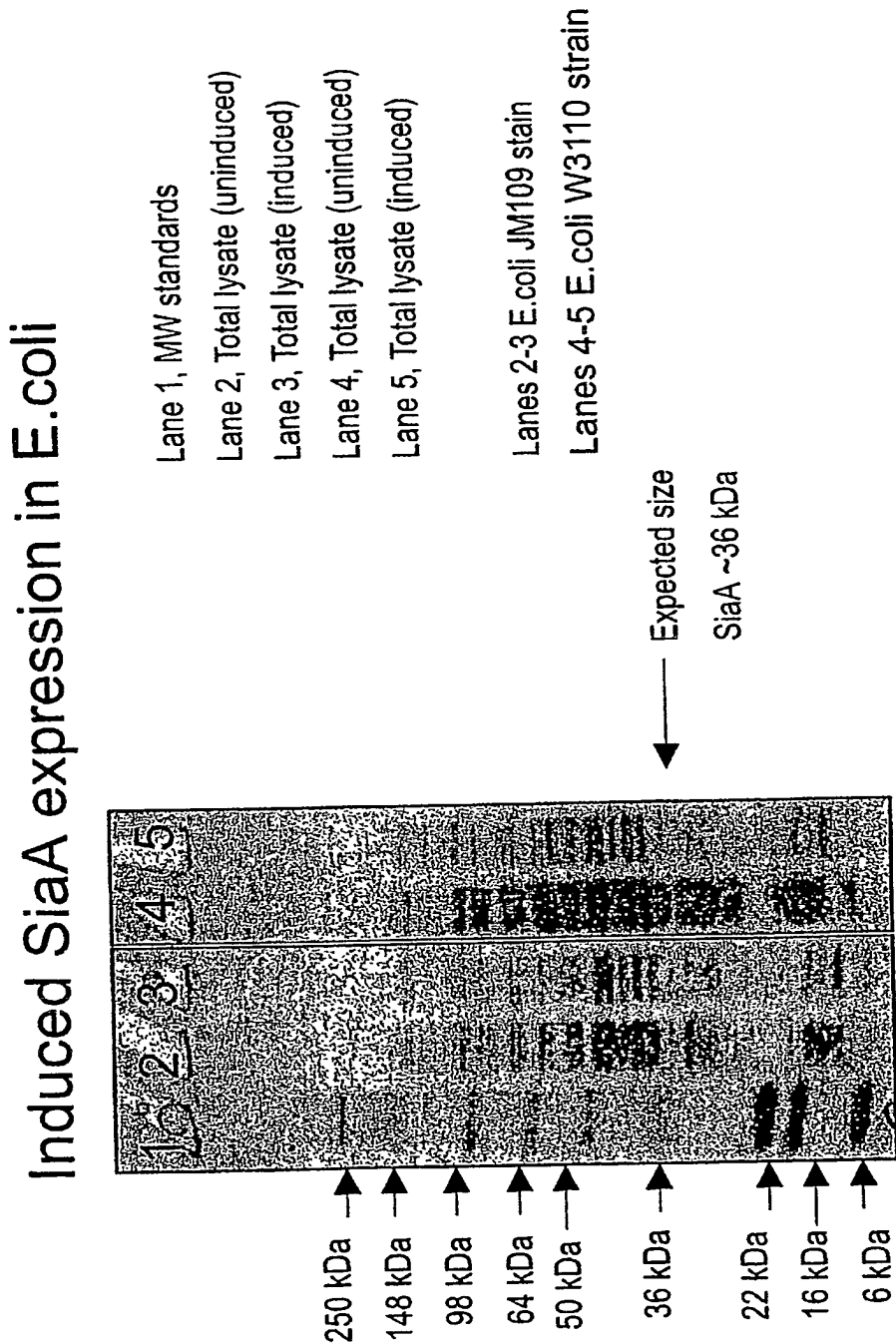SAYGFITEGHERFSDHYYDTSWKRLIFYINHDFRLERMVWKRLHDEGIIWLYQRPQSDKAKN

FIG. 26B

MKIEEGKLVIWINGDKGYNGLAEVGKKFEKDTGIKVTVEHPDKLEEKFPQVAATGDGPDIIFWAHDRF
GGYAQSGLLAEITPDKAFQDKLYPFTWDAVRYNGKLIAYPIAVEALSLIYNKDLLPNPPKTWEEIPALD
KELKAKGKSALMFNLQEPYFTWPLIAADGGYAFKYENGKYDIKDVGVDNAGAKAGLTFLVDLIKNKH
MNADTDYSIAEAAFNKGETAMTINGPWAWSNIDTSKVNYGVTVLPTFKGQPSKPFVGVLSAGINAASP
NKELAKEFLENYLLTDEGLEAVNKDKPLGAVALKSYEEELAKDPRIAATMENAQKGEIMPNIPQMSAF
WYAVRTAVINAASGRQTVDEALKDAQTNSSSNNNNNNNLGEGRISEFGSKEPQTKPSRHQKIE
NIKERSLQSLAKPKSQAPTRARRITIYAEPVPENNALNTQTQPKAHTGDRGKEANQAPPEEQDK
VPHTAQRAAWKSPEKEKTMVNTLSPRGQDAGMASGRTEAQSWKSQDTKTTQGNGGQTRKLTA
SRITVSEKHQGKAAITAKTILPIPKSQHRMLAPTGAVSTRTRQKGVTTAVIPPKEKKPQATPPPAPFQ
SPTTQRNQRLKAANFKSEPRWDFEEKYSFEIGGLQTTCPDSVKIKASKSLWLQKLFLPNLTLFLDS
RHFNQSEWDRLEHFAPPFGFMELNYSLVQKVVTRFPPVPQQQLLASLPAGSLRCITCAVVGNG
GILNNSHMGQEIDSHDYVFRLSGALIKGYEQDVGTRISFYGFTAFSLTQSLLILGNRGFKNVPLGK
DVRYLHFLEGTRDYEWLEALLMNQTVMSKNLFWFRHRPQEAFREALHMDRYLLLHPDFLRYM
KNRFLRSKTLDGAHWRIYRPTTGALLLTALQLCDQVSAYGFITEGHERFSDHYYDTSWKRLIFY
INHDFKLEREVWKRLHDEGIIRLYQRPGPGTAKAKN

Kda

Lanes: 1) See Blue Plus 2
2) MBP-GalNAcT2
3) MBP-GalNAcT2 (conc)
4) Co-refold
5) Co-refold (conc)
6) MBP-ST3Gal1
7) MBP-ST3Gal1 (conc)

250 —

148 —

98 —　　　　　　　　　　　　　← MBP-GalNAcT2
　　　　　　　　　　　　　　　← MBP-ST3Gal1
64 —

50 —

36 —

22 —

FIG. 27A

Kda

Lanes:
1) See Blue Plus 2
2) Control enzymes 0hrs
3) Control enzymes 16hrs

4) Separate refold 0hrs
5) Separate refold 16hrs
6) Co-refold 0hrs
7) Co-refolded 16hrs

250 —
148 —　　　　　　　　　　　← MBP-GalNAcT2
98 —　　　　　　　　　　　　← MBP-ST3Gal1
64 —　　　　　　　　　　　　← IFα-2b-20K PEG
50 —

36 —

22 —
14 —　　　　　　　　　　　　← IFα-2b

FIG. 27B

Induced SiaA expression in E.coli

Lane 1, MW standards
Lane 2, Total lysate (uninduced)
Lane 3, Total lysate (induced)
Lane 4, Total lysate (uninduced)
Lane 5, Total lysate (induced)

Lanes 2-3 E.coli JM109 stain
Lanes 4-5 E.coli W3110 strain

Expected size
SiaA ~36 kDa

250 kDa
148 kDa
98 kDa
64 kDa
50 kDa
36 kDa
22 kDa
16 kDa
6 kDa

There is no obvious inducible band at the expected

Mass of 36 kDa for the native SiaA protein.

FIG. 28

Induced SiaA/MBP expression in E.coli

Expected size

SiaA/MBP ~86 kDa

Lane 1, MW standards

Lane 2, Total lysate (uninduced) colony A

Lane 3, Total lysate (induced) colony A

Lane 4, Total lysate (uniduced) colony B

Lane 5, Total lysate (induced) colony B

Lane 6, Total lysate (induced) colony C

Lanes 2-6 E.coli JM109 stain

250 kDa
148 kDa
98 kDa
64 kDa
50 kDa
36 kDa
22 kDa
16 kDa
6 kDa
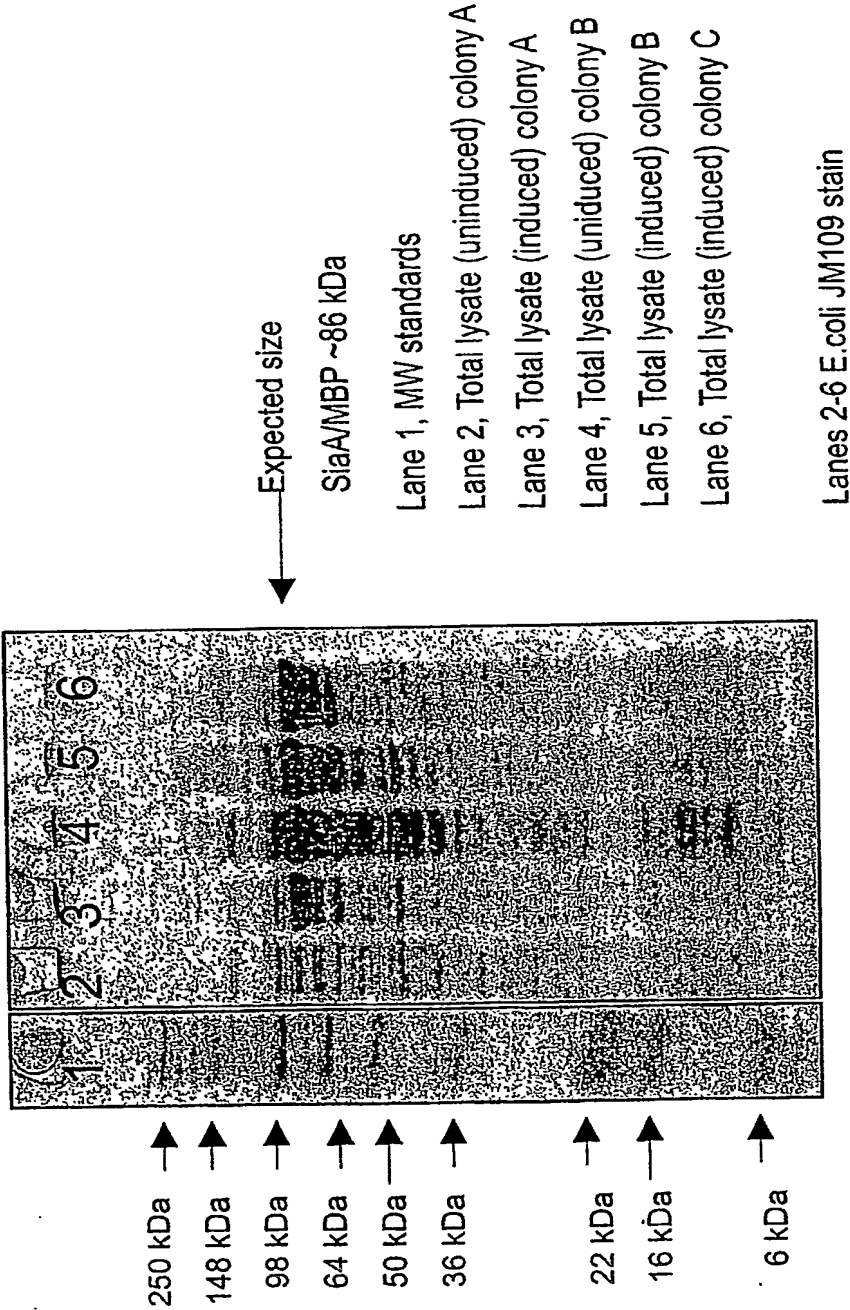
High level production of SiaA/MBP even in absence of IPTG induction.
Compare to figure X where SiaA production is not obvious. The presence of
the fusion partner (MBP) drives high levels of expression.

FIG. 29

```
  1 mkfrepllgg saampgaslq racrllvavc alhlgvtlvy ylagrdlrrl pqlvgvhppl
 61 qgsshgaaai gqpsgelrlr gvappplqn  sskprsraps nldayshpgp gpgpgsnlts
121 apvpstttrs ltacpeespl lvgpmliefn ipvdlklieq qnpkvklggr ytpmdcisph
181 kvaiiilfrn rqehlkywly ylhpmvqrqq ldygiyving agesmfnrak llnvgfkeal
241 kdydyncfvf sdvdlipmnd hntyrcfsqp rhisvamdkf gfslpyvqyf ggvsalskqq
301 flsingfpnn ywgwggeddd iynrlafrgm svsrpnavig kcrmirhsrd kknepnpqrf
361 driahtketm lsdglnslty mvlevqrypl ytkitvdigt ps
```

FIG. 30

Bovine (b) GalT1 (β1,4GalT1) constructs
UDP-galactose β-N-acetylglucosaminide β4-galactosyltransferase (EC 2.4.1.38)

FIG. 31

37/54
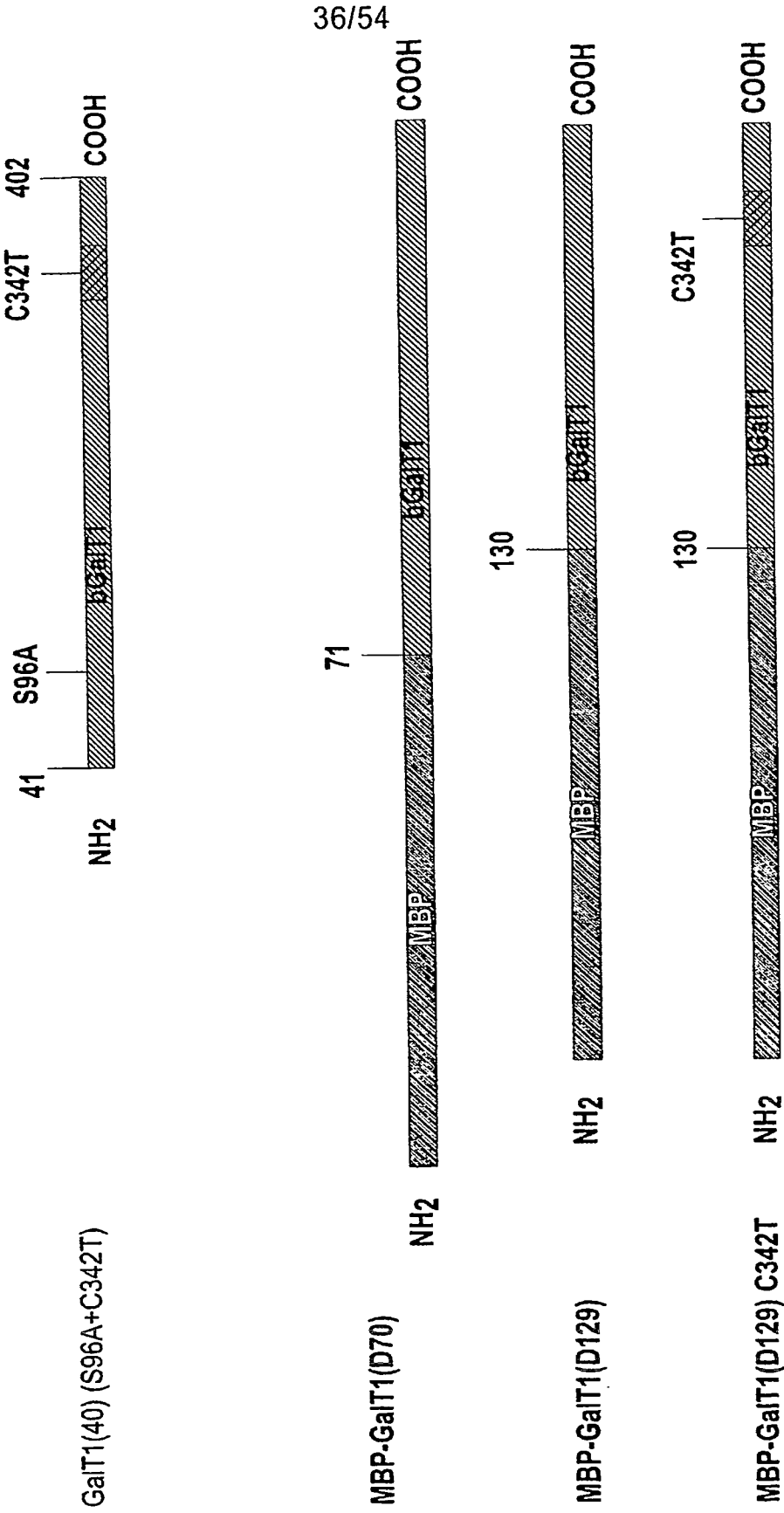
GalT1 TLC assay
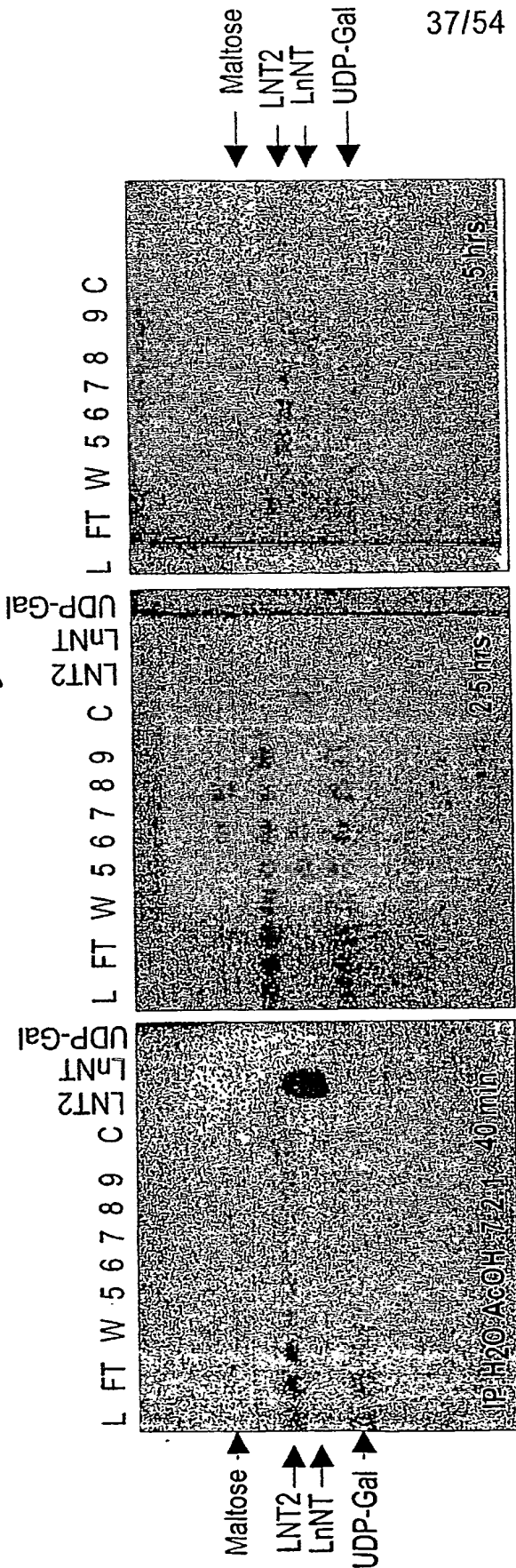
LnNT reactions with

L: Loaded sample (refolded, dialyzed MBP-GalT1(70) onto amylose column)
FT: Amylose column Flow trough
W: Amylose column wash
5: Maltose Eluted Fr # 5
6: Maltose Eluted Fr # 6
7: Maltose Eluted Fr # 7
8: Maltose Eluted Fr # 8
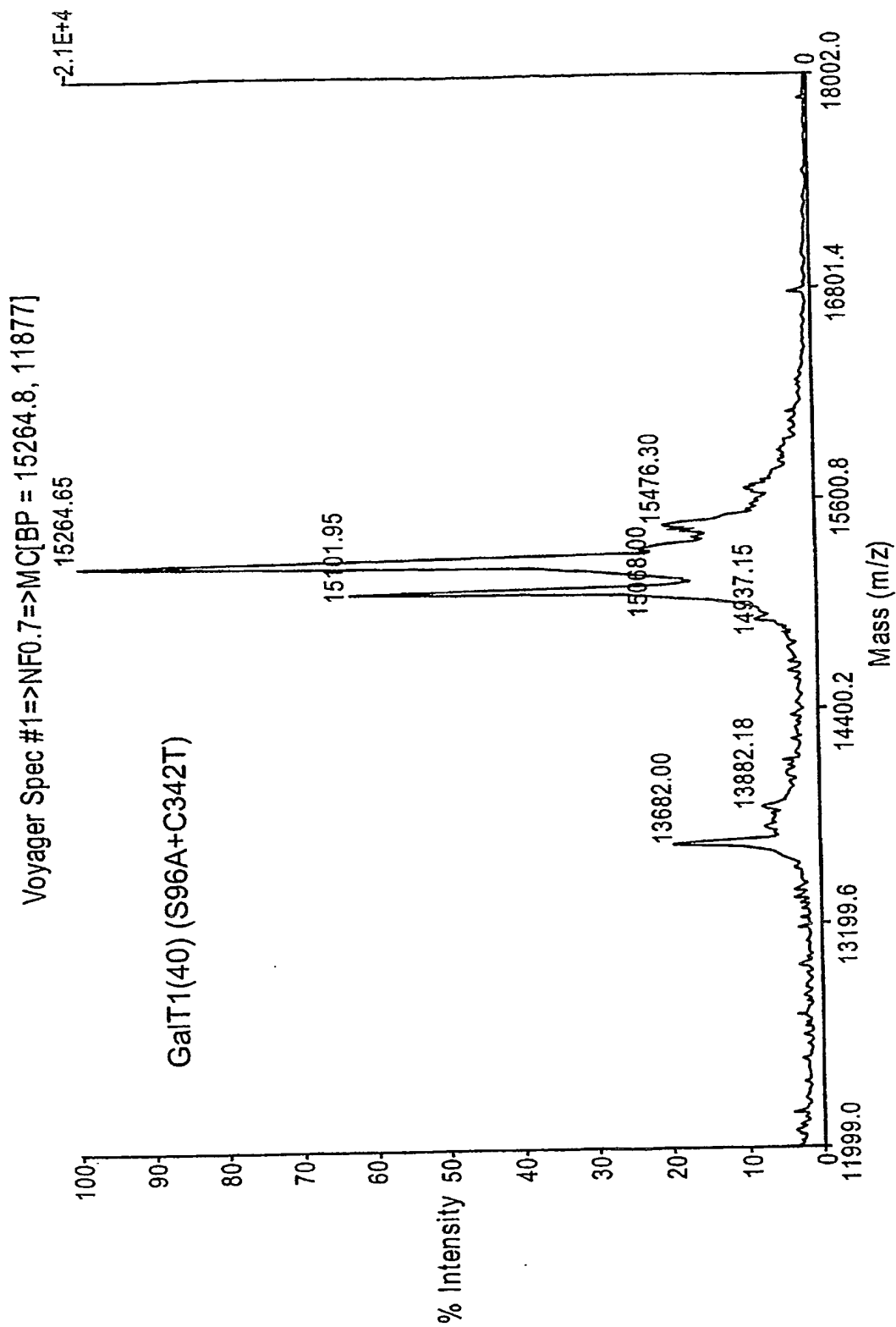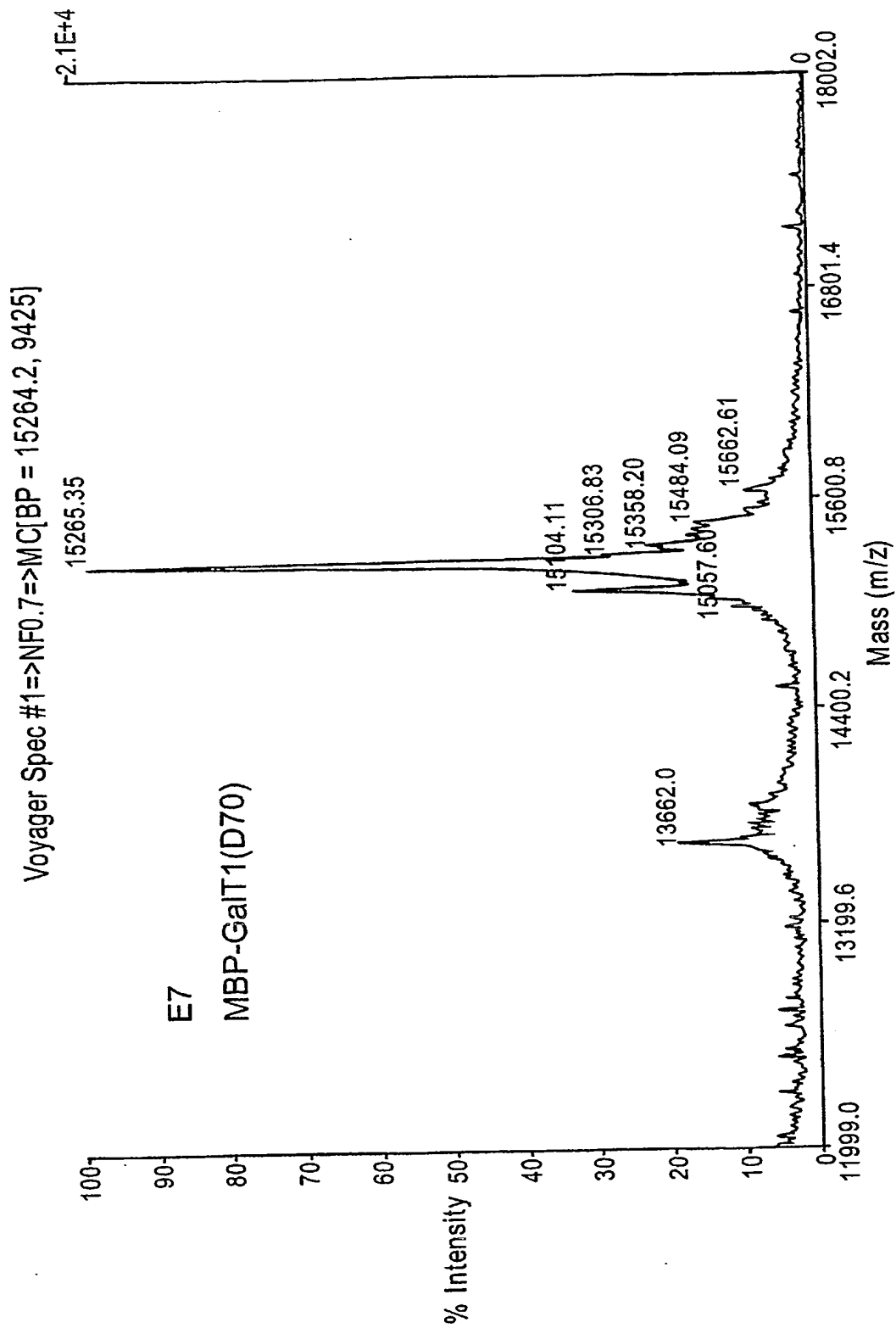9: Maltose Eluted Fr # 9
C: Control with water

FIG. 32

FIG. 33

FIG. 33 (CONT.)

FIG. 33 (CONT.)

Voyager Spec #1=>NF0.7=>MC[BP = 15263.2, 7633]

E8
MBP-GalT1(D70)
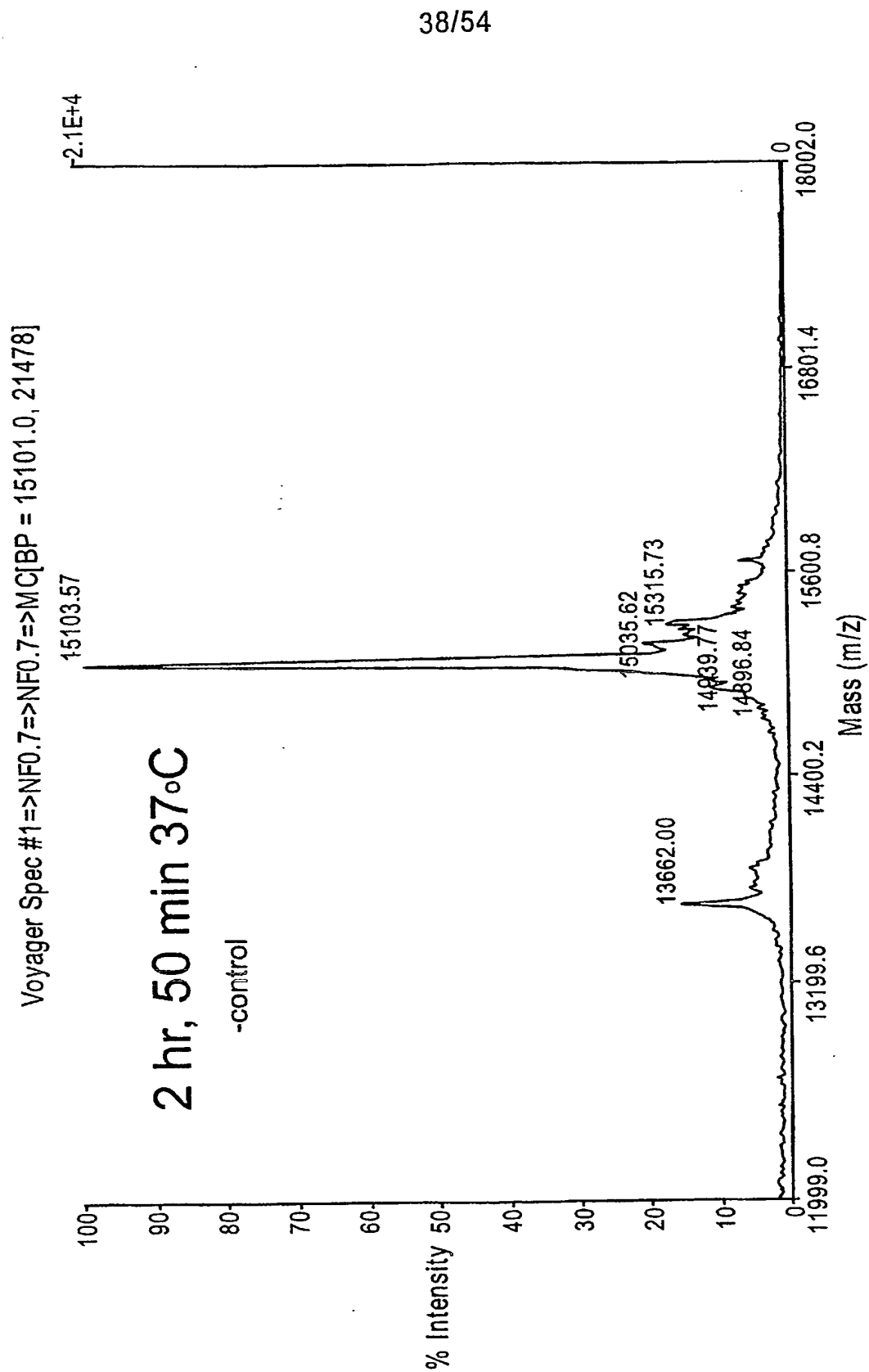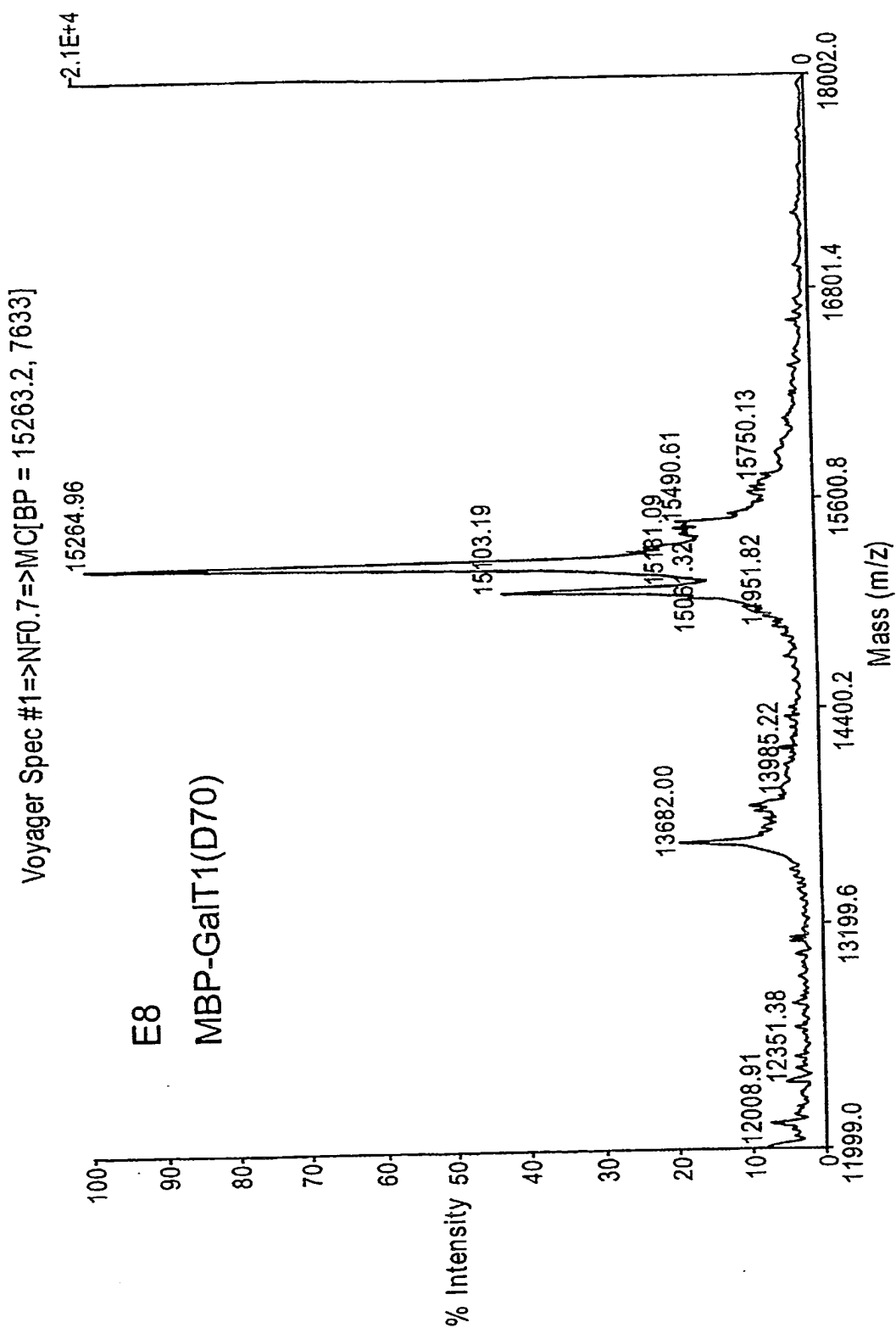
FIG. 33 (CONT.)

Kinetics of RNase B modelling with GalT1 6.5.03

FIG. 34

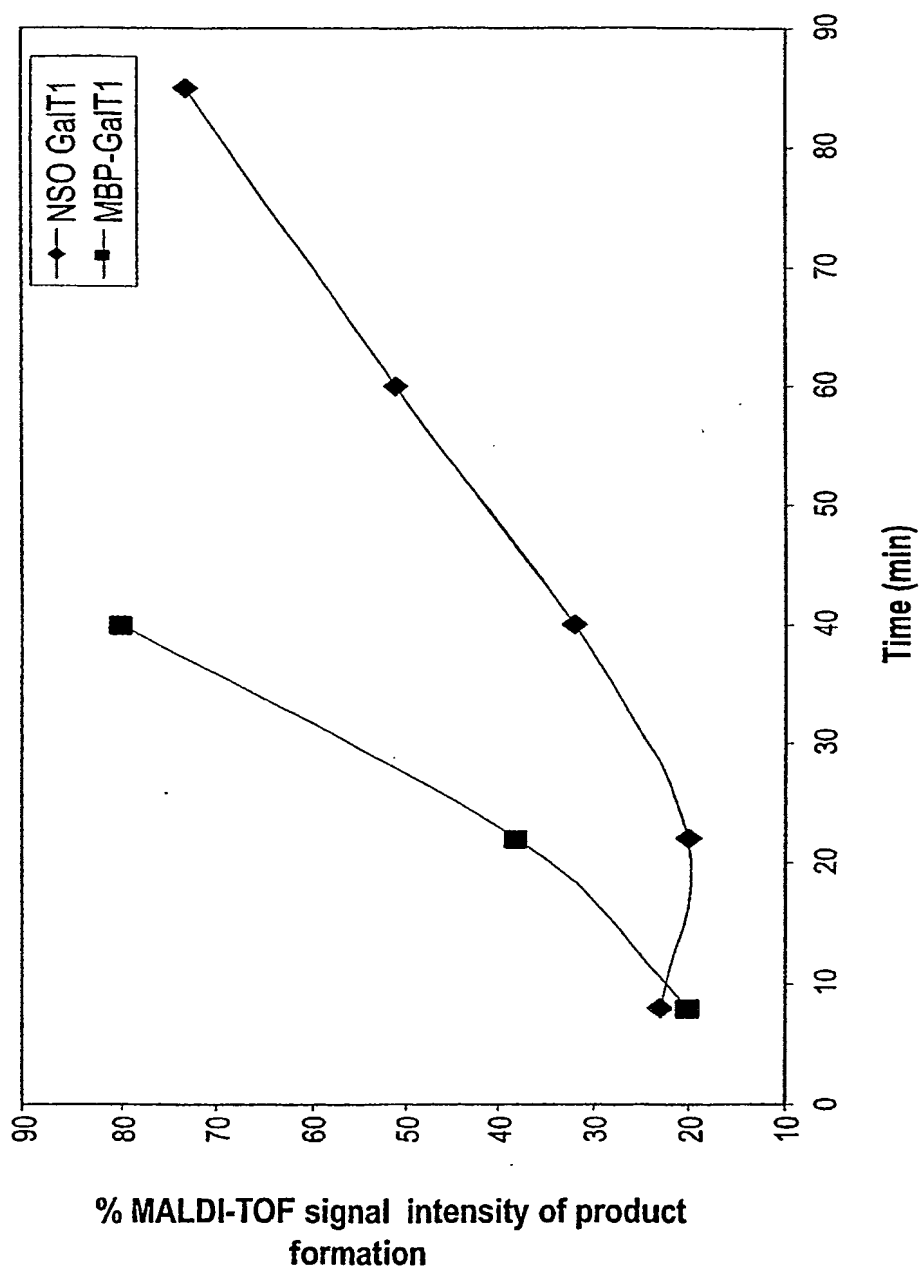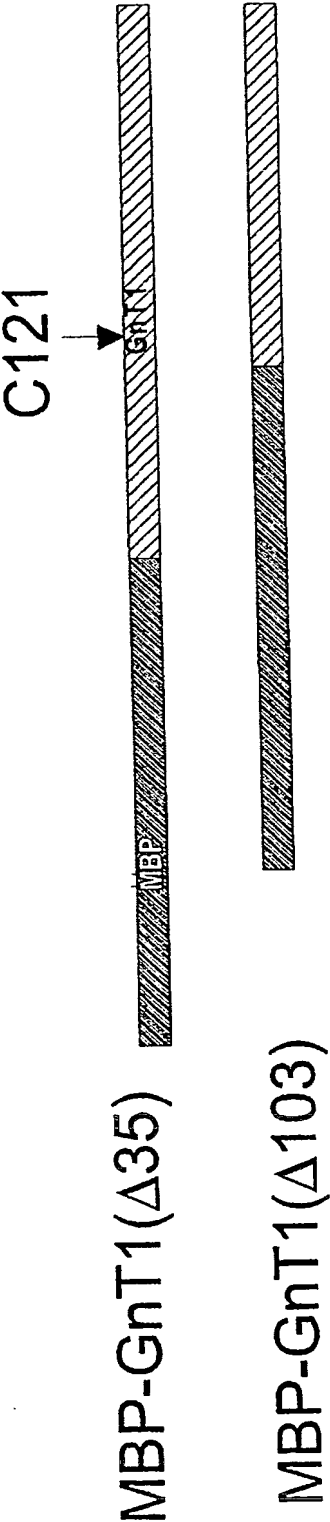GnT1 constructs

C121

GnT1

MBP

MBP-GnT1(Δ35)

MBP-GnT1(Δ103)

```
1    mlkkqsaglv  lwgailfvaw  nallllffwt  rpapgrppsv  saldgdpasl  trevirlaqd
61   aevelerqrg  llqqigdals  sqrgrvptaa  ppaqprvpvt  papavipilv  iacdrstvrr
121  cldkllhyrp  saelfpiivs  qdcgheetaq  aiasygsavt  hirqpdlssi  avppdhrkfq
181  gyykiarhyr  walgqvfrqf  rfpaavvved  dlevapdffe  yfratypllk  adpslwcvsa
241  wndngkeqmv  dasrpellyr  tdffpglgwl  llaelwaele  pkwpkafwdd  wmrrpeqrgg
301  racirpeisr  tmtfgrkgvs  hgqffdqhlk  fiklnqqfvh  ftqldlsylq  reaydrdfla
361  rvygapqlqv  ekvrtndrke  lgevrvqytg  rdsfkafaka  lgvmddlksg  vpragyrgiv
421  tfqfrgrrvh  lappptwegy  dpswn
```

FIG. 35

Comparison of refolded Δ103 and Δ35 truncations of MBP-GnT1

MBP-GnT1(Δ35)
MBP-GnT1(Δ103)

1- MBP-GnT1(Δ35) C121A
2- MBP-GnT1(Δ103) R120A + C121H
3- MBP-GnT1(Δ103) C121A

SS 636-49

Activities of Two different batches (A1 and A2) of Refolded* MBP-GnT1(Δ35) C121A at different time points
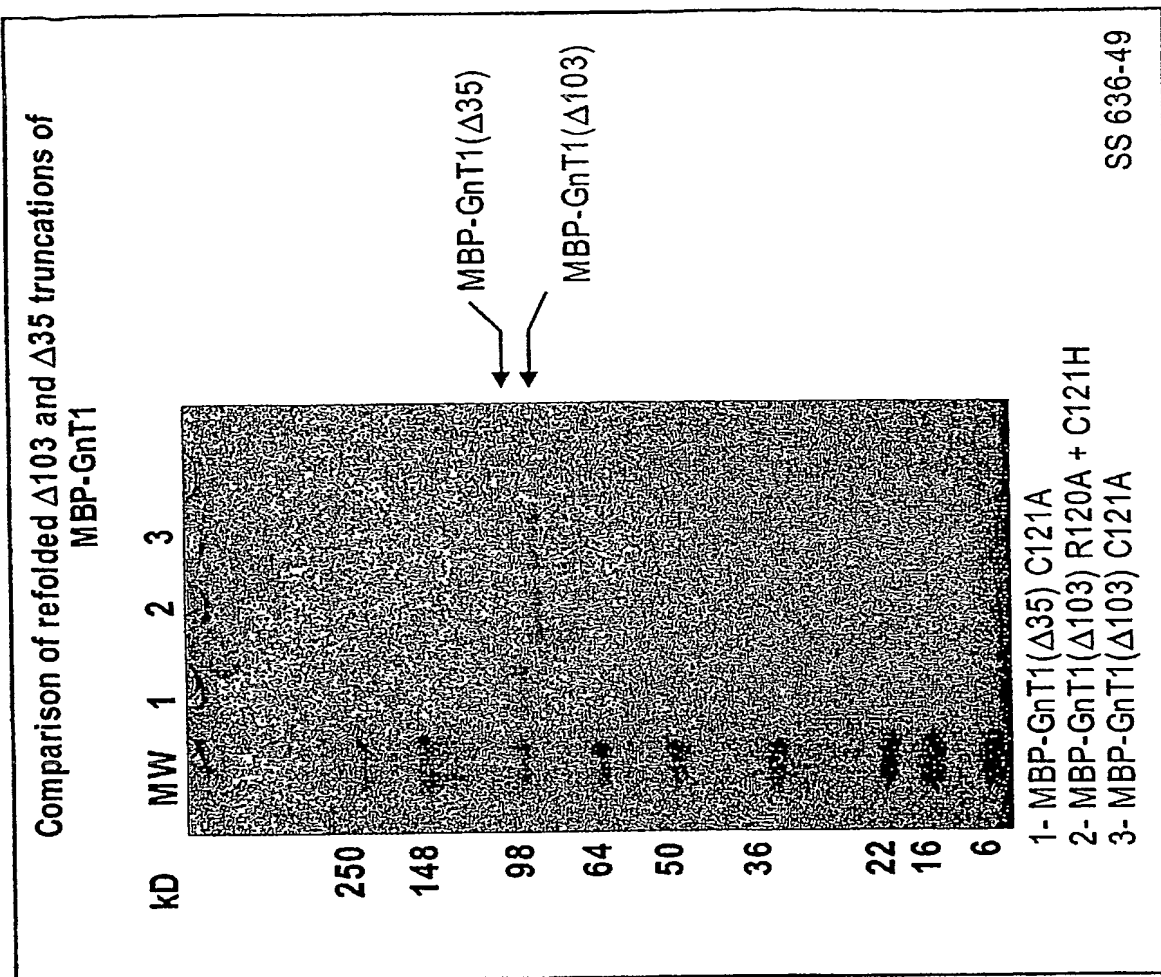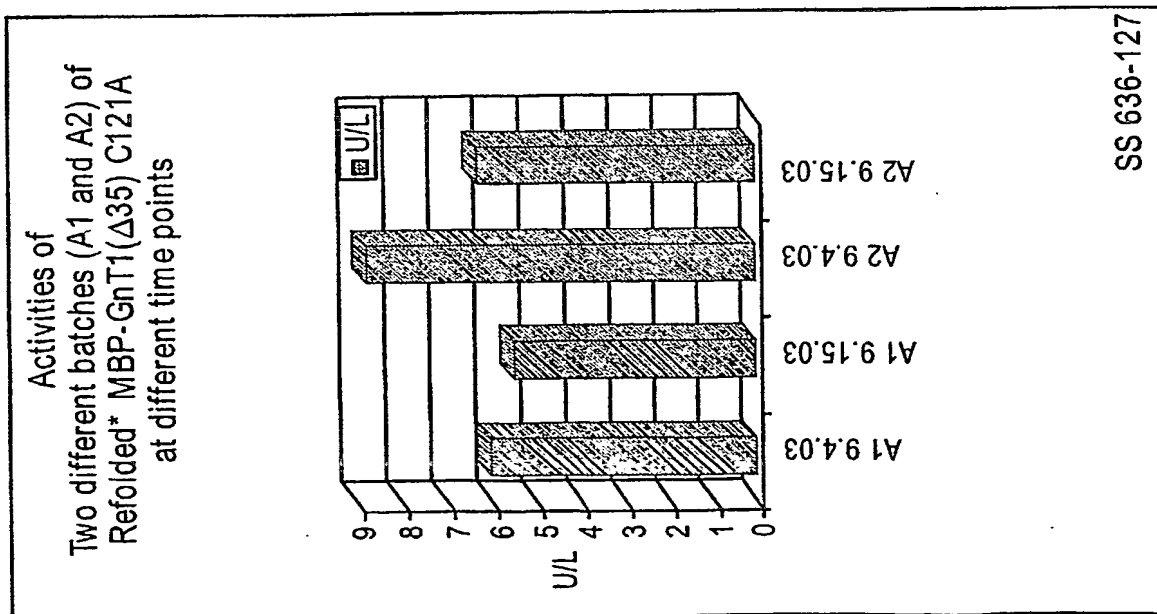
SS 636-127

FIG. 36

```
  1 mapmrkkstl klitllvlfi fltsfflnys htvvttawfp kqmvielsen fkklmkypyr
 61 pctctrciee qrvsawfder fnrsmqpllt aknahleedt ykwwlrlqre kqpnnlndti
121 relfqvvpgn vdpllekrlv scrrcavvgn sgnlkesyyg pqidshdfvl rmnkaptegf
181 eadvgsktth hfvypesfre laqevsmilv pfkttdlewv isatttgtis htyvpvpaki
241 kvkkekiliy hpafikyvfd rwlqghgryp stgilsvifs lhicdevdly gfgadskgnw
301 hhywennpsa gafrktgvhd gdfesnvtti lasinkirif kgr
```

FIG. 37

Human ST6GalNAcI
MRSCLWRCRHLSQGVQWSLLLAVLVFFLFALPSFIKEPQTKPSRHQRTENIKERSLQS
LAKPKSQAPTRARRTTIYAEPVPENNALNTQTQPKAHTTGDRGKEANQAPPEEQDK
VPHTAQRAAWKSPEKEKTMVNTLSPRGQDAGMASGRTEAQSWKSQDTKTTQGNG
GQTRKLTASRTVSEKHQGKAATTAKTLIPKSQHRMLAPTGAVSTRTRQKGVTTAVIP
PKEKKPQATPPPAPFQSPTTQRNQRLKAANFKSEPRWDFEEKYSFEIGGLQTTCPDSV
KIKASKSLWLQKLFLPNLTLFLDSRHFNQSEWDRLEHFAPPFGFMELNYSLVQKVVT
RFPPVPQQQLLLASLPAGSLRCITCAVVGNGGILNNSHMGQEIDSHDYVFRLSGALIK
GYEQDVGTRTSFYGFTAFSLTQSLLILGNRGFKNVPLGKDVRYLHFLEGTRDYEWLE
ALLMNQTVMSKNLFWFRHRPQEAFREALHMDRYLLLHPDFLRYMKNRFLRSKTLD
GAHWRIYRPTTGALLLLTALQLCDQVSAYGFITEGHERFSDHYYDTSWKRLIFYINH
DFKLEREVWKRLHDEGIIRLYQRPGPGTAKAKN

FIG. 38A

Chicken ST6GalNAcI
MGFLIRRLPKDSRIFRWLLILTVFSFIITSFSALFGMEKSIFRQLKIYQSIAHMLQVDTQ
DQQGSNYSANGRISKVGLERDIAWLELNTAVSTPSGEGKEEQKKTVKPVAKVEEAK
EKVTVKPFPEVMGITNTTASTASVVERTKEKTTARPVPGVGEADGKRTTIALPSMKE
DKEKATVKPSFGMKVAHANSTSKDKPKAEEPPASVKAIRPVTQAATVTEKKKLRAA
DFKTEPQWDFDDEYILDSSSPVSTCSESVRAKAAKSDWLRDLFLPNITLFIDKSYFNV
SEWDRLEHFAPPYGFMELNYSLVEEVMSRLPPNPHQQLLLANSSSNVSTCISCAVVG
NGGILNNSGMGQEIDSHDYVFRVSGAVIKGYEKDVGTKTSFYGFTAYSLVSSLQNLG
HKGFKKIPQGKHIRYIHFLEAVRDYEWLKALLLDKDIRKGFLNYYGRRPRERFDEDF
TMNKYLVAHPDFLRYLKNRFLKSKNLQKPYWRLYRPTTGALLLLTALHLCDRVSAY
GYITEGHQKYSDHYYDKEWKRLVFYVNHDFNLEKQVWKRLHDENIMKLYQRS

FIG. 38B

Mouse ST6GalNAcI protein beginning at residue 32 of the native mouse protein
DPRAKDSRCQFIWKNDASAQENQQKAEPQVPIMTLSPRVHNKESTSVSSKDLKKQER
EAVQGEQAEGKEKRKLETIRPAPENPQSKAEPAAKTPVSEHLDKLPRTPGALSTRKTP
MATGAVPAKKKVVQATKSPASSPHPTTRRRQRLKASEFKSEPRWDFEEEYSLDMSSL
QTNCSASVKIKASKSPWLQNIFLPNITLFLDSGRFTQSEWNRLEHFAPPFGFMELNQSL
VQKVVTRFPPVRQQQLLLASLPTGYSKCITCAVVGNGGILNDSRVGREIDSHDYVFR
LSGAVIKGYEQDVGTRTSFYGFTAFSLTQSILILGRRGFQHVPLGKDVRYLHFLEGTR
NYEWLEAMFLNQTLAKTHLSWFRHRPQEAFRNALDLDRYLLLHPDFLRYMKNRFL
RSKTLDTAHWRIYRPTTGALLLLTALHLCDKVSAYGFITEGHQRFSDHYYDTSWKRL
IFYINHDFRLERMVWKRLHDEGIIWLYQRPQSDKAKN
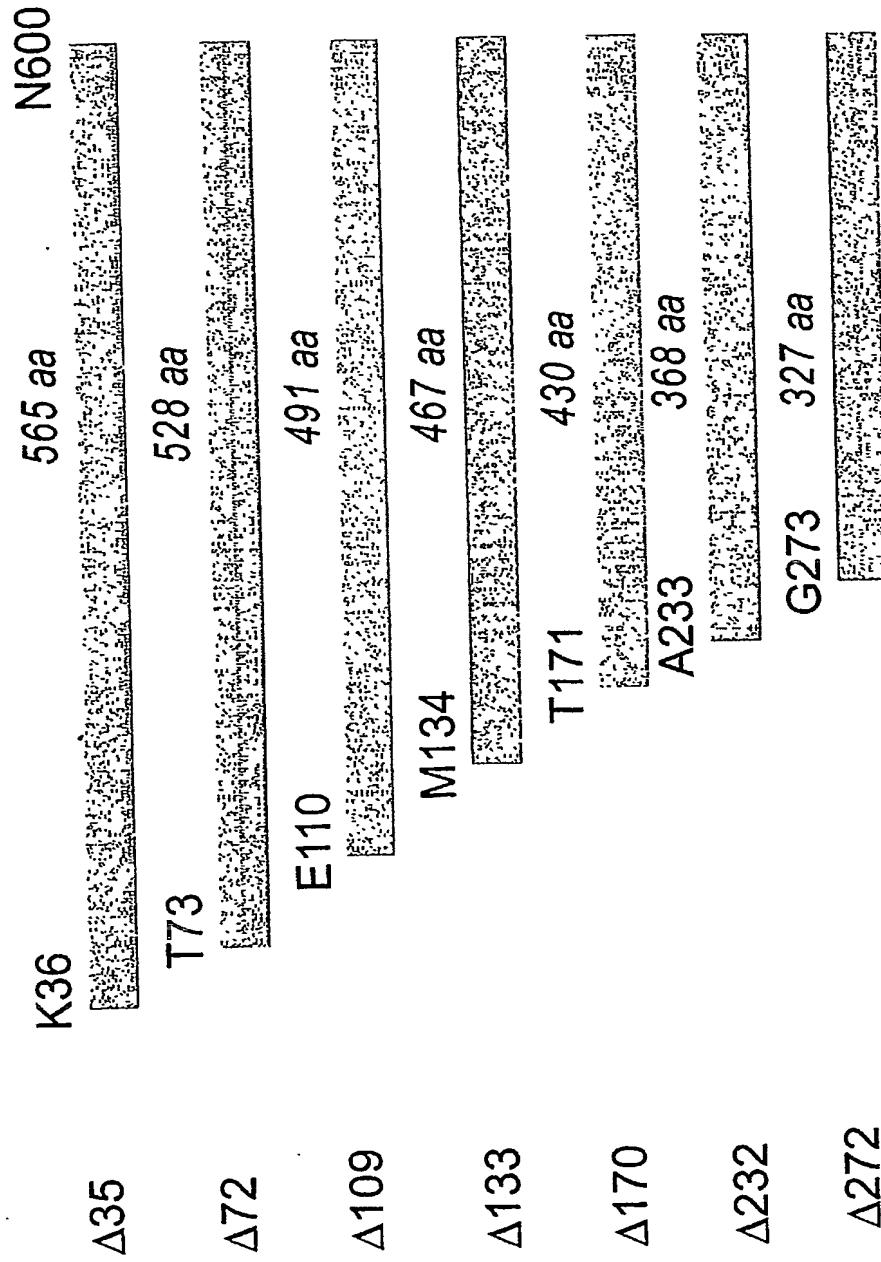
FIG. 38C

Hum (h)ST6GalNAcI truncations

FIG. 39

MBP-hST6GalNAcI constructs

MBP-hST6GalNAcI (Δ35), 107.6 kD

MBP-hST6GalNAcI (Δ72), 103.3 kD

MBP-hST6GalNAcI (Δ109), 99.2 kD

MBP-hST6GalNAcI (Δ133), 96.7 kD

MBP-hST6GalNAcI (Δ170), 92.7 kD

MBP-hST6GalNAcI (Δ232), 86.0 kD

MBP-hST6GalNAcI (Δ272), 81.4 kD

K36

T73

E110

M134

T171

A233

G273

MBP   ST6GalNAcI

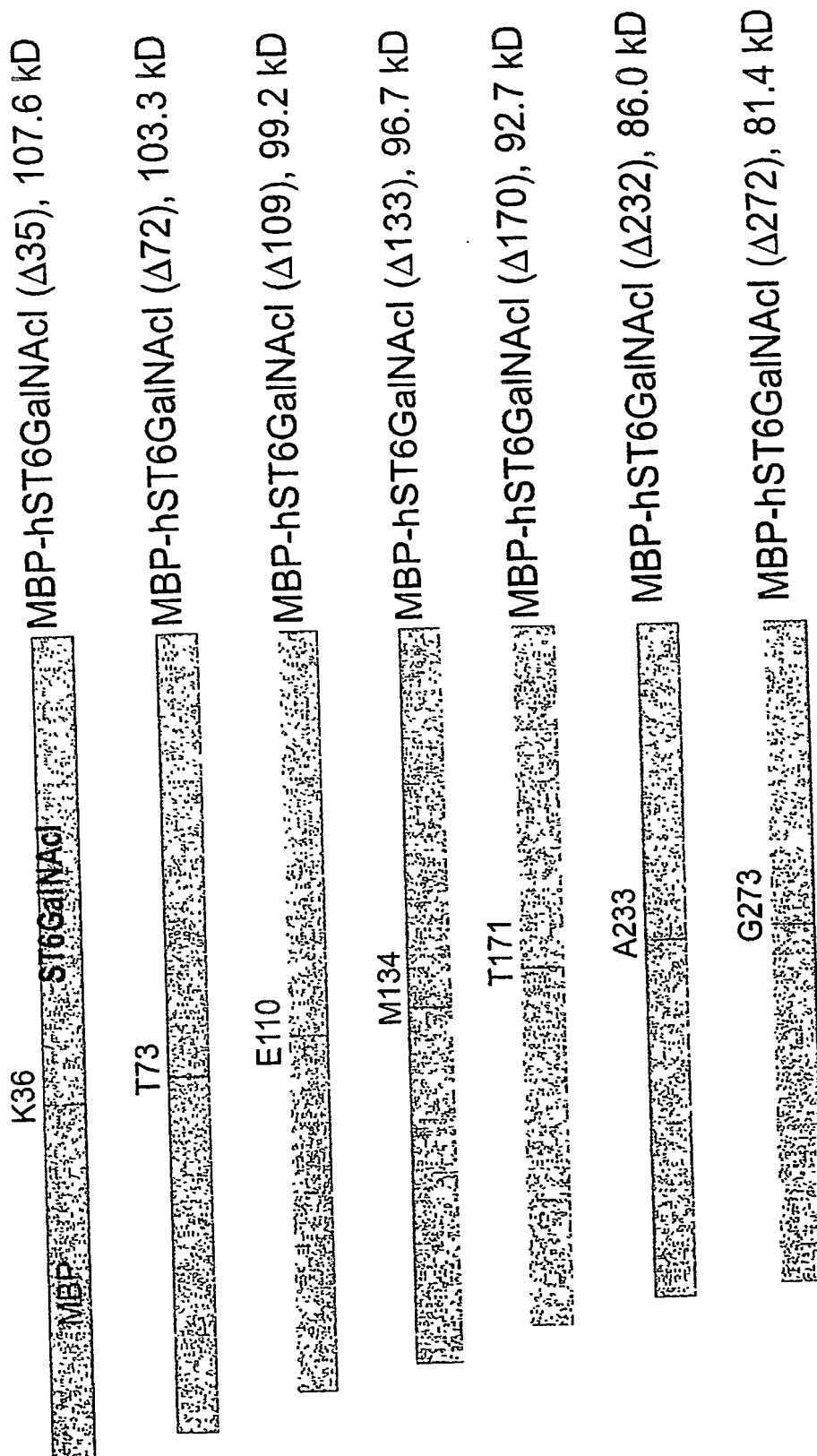FIG. 40

MASKSWLNFLTFLCGSAIGFLLCSQLFSILLGEKVDTQPNVLHNDPHARHSDDNGQN
HLEGQMNFNADSSQHKDENTDIAENLYQKVRILCWVMTGPQNLEKKAKHVKATW
AQRCNKVLFMSSEENKDFPAVGLKTKEGRDQLYWKTIKAFQYVHEHYLEDADWFL
KADDDTYVILDNLRWLLSKYDPEEPIYFGRRFKPYVKQGYMSGGAGYVLSKEALKR
FVDAFKTDKCTHSSSIEDLALGRCMEIMNVEAGDSRDTIGKETFHPFVPEHHLIKGYL
PRTFWYWNYNYYPPVEGPGCCSDLAVSFHYVDSTTMYELEYLVYHLRPYGYLYRY
QPTLPERILKEISQANKNEDTKVKLGNP

FIG. 41

Sequence	Size	Range	Mode					
SC1GALT1.AMI	342	1-	342 Normal					
SPTS122.AMI	342	1-	342 Normal					

```
                      10          20          30          40          50
SC1GALT1.AMI    EFMPYDGHRH  GDVNDAHHSH  DMMEMSGPEQ  DVGGHEHVHE  NSTIAERLYS
SPTS122.AMI     EFMPYDGHRH  GDVNDAHHSH  DMMEMSGPEQ  DVGGHEHVHE  NSTIAERLYS

                      60          70          80          90         100
SC1GALT1.AMI    EVRVLCWIMT  NPSNHQKKAR  HVKRTWGKRC  NKLIFMSSAK  DDELDAVALP
SPTS122.AMI     EVRVLCWIMT  NPSNHQKKAR  HVKRTWGKRC  NKLIFMSSAK  DDELDAVALP

                     110         120         130         140         150
SC1GALT1.AMI    VGEGRNNLWG  KTKEAYKYIY  EHHINDADWF  LKADDDTYTI  VENMRYMLYP
SPTS122.AMI     VGEGRNNLWG  KTKEAYKYIY  EHHINDADWF  LKADDDTYTI  VENMRYMLYP

                     160         170         180         190         200
SC1GALT1.AMI    YSPETPVYFG  CKFKPYVKQG  YMSGGAGYVL  SREAVRRFVV  EALPNPKLCK
SPTS122.AMI     YSPETPVYFG  CKFKPYVKQG  YMSGGAGYVL  SREAVRRFVV  EALPNPKLCK

                     210         220         230         240         250
SC1GALT1.AMI    SDNSGAEDVE  IGKCLQNVNV  LAGDSRDSNG  RGRFFPFVPE  HHLIPSHTDK
SPTS122.AMI     SDNSGAEDVE  IGKCLQNVNV  LAGDSRDSNG  RGRFFPFVPE  HHLIPSHTDK

                     260         270         280         290         300
SC1GALT1.AMI    KFWYWQYIFY  KTDEGLDCCS  DNAISFHYVS  PNQMYVLDYL  IYHLRPYGII
SPTS122.AMI     KFWYWQYIFY  KTDEGLDCCS  DNAISFHYVS  PNQMYVLDYL  IYHLRPYGII

                     310         320         330         340         350
SC1GALT1.AMI    NTPDALPNKL  AVGELMPEIK  EQATESTSDG  VSKRSAETKI  Q* ......
SPTS122.AMI     NTPDALPNKL  AVGELMPEIK  EQATESTSDG  VSKRSTETKI  Q* ......
```

FIG. 42

Yersinia MBP

```
MKIEEGKLVI  WINGDKGYNG  LAEVGKKFEK  DTGIKVTIEH  PDKLEEKFPQ
VAATGDGPDI  IFWAHDRFGG  YAQSGLLAEL  TPSKAFQEKL  FPFTWDAVRF
NGKLIGYPVA  VEALSLIYNK  DLVKEAPKTW  EEIPALDKTL  RANGKSAIMW
NLQEPYFTWP  VIAADGGYAF  KFENGVYDAK  NVGVNNAGAQ  AGLQFIVDLV
KNKHINADTD  YSIAEAAFNK  GETAMTINGP  WAWSNIDKSK  INYGVTLLPT
FHGQPSKPFV  GVLTAGINAA  SPNKELATEF  LENYLITDQG  LAEVNKDKPL
GAVALKSFQE  QLAKDPRIAA  TMDNATNGEI  MPNIPQMAAF  WYATRSAVLN
AITGRQTVEA  ALNDAATRIT  K
```

FIG. 43A

E. coli MBP

```
MKIEEGKLVI  WINGDKGYNG  LAEVGKKFEK  DTGIKVTVEH  PDKLEEKFPQ
VAATGDGPDI  IFWAHDRFGG  YAQSGLLAEI  TPDKAFQDKL  YPFTWDAVRY
NGKLIAYPIA  VEALSLIYNK  DLLPNPPKTW  EEIPALDKEL  KAKGKSALMF
NLQEPYFTWP  LIAADGGYAF  KYENGKYDIK  DVGVDNAGAK  AGLTFLVDLI
KNKHMNADTD  YSIAEAAFNK  GETAMTINGP  WAWSNIDTSK  VNYGVTVLPT
FKGQPSKPFV  GVLSAGINAA  SPNKELAKEF  LENYLLTDEG  LEAVNKDKPL
GAVALKSYEE  ELAKDPRIAA  TMENAQKGEI  MPNIPQMSAF  WYAVRTAVIN
AASGRQTVDE  ALKDAQTNS
```

FIG. 43B

Pyrococcus furiosus MBP

```
MKIEEGKVVI  WHAMQPNELE  VFQSLAEEYM  ALSPEVEIVF  EQKPNLEDAL
KAAIPTGQGP  DLFIWAHDWI  GKFAEAGLLE  PIDEYVTEDL  LNEFAPMAQD
AMQYKGHYYA  LPFAAETVAI  IYNKEMVSEP  PKTFDEMKAI  MEKYYDPANE
KYGIAWPINA  YFISAIAQAF  GGYYFDDKTE  QPGLDKPETI  EGFKFFFTEI
WPYMAPTGDY  NTQQSIFLEG  RAPMMVNGPW  SINDVKKAGI  NFGVVPLPPI
IKDGKEYWPR  PYGGVKLIYF  AAGIKNKDAA  WKFAKWLTTS  EESIKTLALE
LGYIPVLTKV  LDDPEIKNDP  VIYGFGQAVQ  HAYLMPKSPK  MSAVWGGVDG
AINEILQDPQ  NADIEGILKK  YQQEILNNMQ  G
```

FIG. 43C

Thermococcus litoralis MBP

```
MKIEEGKIVF  AVGGAPNEIE  YWKGVIAEFE  KKYPGVTVEL  KRQATDTEQR
RLDLVNALRG  KSSDPDVFLM  DVAWLGQFIA  SGWLEPLDDY  VQKDNYDLSV
FFQSVINLAD  KQGGKLYALP  VYIDAGLLYY  RKDLLEKYGY  SKPPETWQEL
VEMAQKIQSG  ERETNPNFWG  FVWQGKQYEG  LVCDFVEYVY  SNGGSLGEFK
DGKWVPTLNK  PENVEALQFM  VDLIHKYKIS  PPNTYTEMTE  EPVRLMFQQG
NAAFERNWPY  AWGLHNADDS  PVKGKVGVAP  LPHFPGHKSA  ATLGGWHIGI
SKYSDNKALA  WEFVKFVESY  SVQKGFAMNL  GWNPGRVDVY  DDPAVVSKSP
HLKELRAVFE  NAVPRPIVPY  YPQLSEIIQK  YVNSALAGKI  SPQEALDKAQ
KEAEELVKQY  SK
```

FIG. 43D

Thermatoga maritime MBP

```
MKIEQTKLTI  WSSEKQVDIL  QKLGEEFKAK  YGIPVEVQYV  DFGSIKSKFL
TAAPQGQGAD  IIVGAHDWVG  ELAVNGLIEP  IPNFSDLKNF  YDTALKAFSY
GGKLYGVPYA  MEAVALIYNK  DYVDSVPKTM  DELIEKAKQI  DEEYGGEVRG
FIYDVANFYF  SAPFILGYGG  YVFKETPQGL  DVTDIGLANE  GAVKGAKLIK
RMIDEGVLTP  GDNYGTMDSM  FKEGLAAMII  NGLWAIKSYK  DAGINYGVAP
IPELEPGVPA  KPFVGVQGFM  INAKSPNKVI  AMEFLTNFIA  RKETMYKIYL
ADPRLPARKD  VLELVKDNPD  VVAFTQSASM  GTPMPNVPEM  APVWSAMGDA
LSIIINGQAS  VEDALKEAVD  KIKAQIEK
```

FIG. 43E

Vibrio cholerae MBP
```
MKIEEGQLTI  WINGDKGYNG  LAEVGKKFEA  DTGIKVTVAH  PDALQDKFPQ
TAATGDGPDI  VFWAHDRFGG  YAEAGLLVEI  KPSAKIQEGI  VDFAWDAVKY
NGKIIGYPIA  VESLSLIYNK  DLVPNPPKSW  EEVAELDAKL  KKEGKSAIMW
NLKEPYFTWP  LMAADGGYAF  KYGVDGYDVK  DAGINNKGVK  DAMNFVKGLV
DKGVISPDMD  YSVSESAFNQ  GNTAMTINGP  WSWGNIEKSG  INYGVTTLPK
FNGQASKPFV  GVLTAGISTA  SPNKDLAVEF  IENYLLTNDG  LRMVNNDKPL
GAVALNSFQR  ELDADARIAA  TMDNAMNGEI  MPNIPQMNAF  WSSAKNAIIN
IVDGRQTVDA  ALADAEKQMT  KP
```

FIG. 43F

53/54

```
                     1                                                50
HSGALNAT1.pep   (1)MRKFAYCKVVLATSLIWVLLDMFLLLYFS----------------ECNKC
HSGALNAT2.pep   (1)MRRRS--RMLLCFAFLWVLGIAYYMYSGGGSALAGGAGGGAGRKEDWNEI
Consensus       (1)MRK A KMLL A IWVL   F L                       D N
                     51                                               100
HSGALNAT1.pep  (35)DEKKERGLPAGDVLEPVQKPHEGP-G------EMGKPVVIPKEDQEKMKFM
HSGALNAT2.pep  (49)DPIKKKDLHHSNGEEKAQSMETLPPGKVRWPDFNQEAYVGGTMVRSGQDP
Consensus      (51)D KK L    E Q    P G    D      I            D
                     101                                              150
HSGALNAT1.pep  (79)FKINQFNLMASEMIALNRSLPDVRLEGCKTKVYPDNLPTTSVVIVFHNEA
HSGALNAT2.pep  (99)YARNKFNQVESDKLRMDRAIPDTRHDQCQRKQWRVDLPATSVVITFHNEA
Consensus     (101)F N FN M SD I L RAIPD R D C K W   LP TSVVI FHNEA
                     151                                              200
HSGALNAT1.pep (129)WSTLLRTVHSVINRSPRHMIEEIVLVDDASERDFLKRPLESYVKKLKVPV
HSGALNAT2.pep (149)RSALLRTVVSVLKKSPPHLIKEIILVDDYSN------DPEDGALLGKIEKV
Consensus     (151) S LLRTV SVI KSP HLI EIILVDD S         P D L     V
                     201                                              250
HSGALNAT1.pep (179)HVIRMEQRSGLIRARLKGAAVSKGQVITFLDAHCECTVGWLEPLLARIKH
HSGALNAT2.pep (194)RVLRNDRREGLMRSRVRGADAAQAKVLTFLDSHCECNEHWLEPLLERVAE
Consensus     (201) VIR D R GLIRARLKGA A A VITFLDAHCEC   WLEPLL RI
                     251                                              300
HSGALNAT1.pep (229)DRRTVVCPIIDVISDDTFEYMAGSDMTYGGFNWKLNFRWYPVPQREMDRR
HSGALNAT2.pep (244)DRTRVVSPIIDVINMDNFQYVGASADLKGGFDWNLVFKWDYMTPEQRRSR
Consensus     (251)DR VV PIIDVI D F YMAAS    GGF W L FKW M        R
                     301                                              350
HSGALNAT1.pep (279)KGDRTLPVRTPTMAGGLFSIDRDYFQEIGTYDAGMDIWGGENLEISFRIW
HSGALNAT2.pep (294)QGNPVAPIKTPMIAGGLFVMDKFYFEELGKYDMVVMDVWGGENLEISFRVW
Consensus     (301) G    PIKTP IAGGLF IDK YF EIG YD  MDIWGGENLEISFRIW
                     351                                              400
HSGALNAT1.pep (329)QCGGTLEIVTCSHVGHVFRKATPYTFPGGTGQIINKNNRRLAEVWMDEFK
HSGALNAT2.pep (344)QCGGSLEIIPCSRVGHVFRKQHPYTFPGGSGTVFARNTRRAAEVWMDEYK
Consensus     (351)QCGGSLEII CS VGHVFRK  PYTFPGGSG I  KN RR AEVWMDEFK
                     401                                              450
HSGALNAT1.pep (379)NFFYIISPGVTKVDYGDISSRVGLRHKLQCKPFSWYLENIYPDSQIPRHY
HSGALNAT2.pep (394)NFYYAAVPSARNVPYGNIQSRLELRKKLSCKPFKWYLENVYPELRVPDHQ
Consensus     (401)NFFY   P    V YG I SRL LR KL CKPF WYLENIYPD  IP H
                     451                                              500
HSGALNAT1.pep (429)FSLGEIRNVETNQCLDNMARKENEKVGIFNCHGMGGNQVFSYTANKEIRT
HSGALNAT2.pep (444)DIAFGALQQGTN-CLDTLGHFADGVVGVYECHNAGGNQEWALTKEKSVKH
Consensus     (451)        N TN CLD LA      VGIF CH GGNQ FA T K IK
                     501                                              550
HSGALNAT1.pep (479)DDLCLDVSKLN---GPVIMLKCHHLKGNQLWEYDPVKLTLQHVNSNQCLDK
HSGALNAT2.pep (493)MDLCLTVVDRAPGSLIKLQGCRENDSRQKWEQIEGNSKLRHVGSNLCLDS
Consensus     (501) DLCL V    I L C   Q WE      L HV SN CLD
                     551                              584
HSGALNAT1.pep (527)ATEEDSQVPSIRDCNGSRSQQWLLRNVTLPEIF-
HSGALNAT2.pep (543)RTAK-SGGLSVEVCGPALSQQWKFTLNLQQ----
Consensus     (551) T   S   SI C  A SQQW
```
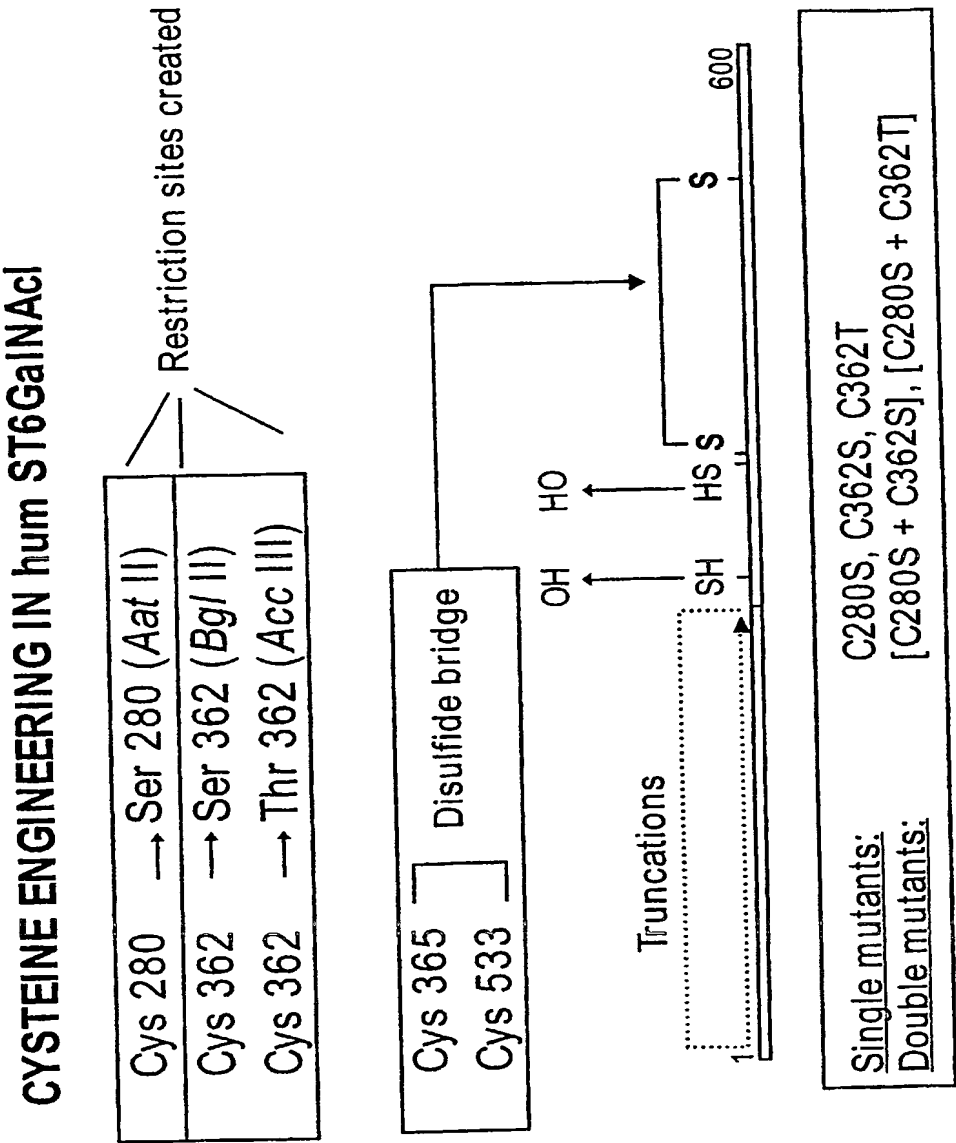
FIG. 44

SUBSTITUTE SHEET (RULE 26)

CYSTEINE ENGINEERING IN hum ST6GalNAcI

Restriction sites created

Cys 280	→Ser 280 (Aat II)
Cys 362	→Ser 362 (Bgl II)
Cys 362	→Thr 362 (Acc III)

Cys 365	Disulfide bridge
Cys 533	

Truncations

OH   OH

SH   HS S

S

1   600

Single mutants: C280S, C362S, C362T
Double mutants: [C280S + C362S], [C280S + C362T]

FIG. 45